

# Selection and Hiring of Quality Police Officers

Prepared for the:

Ohio Office of Criminal Justice Services (OCJS)

September 12, 2008

James Frank, Ph.D.

Billy Henson, MS.

Bradford Reynolds, MS.

Charles Klahm IV, MS.

University of Cincinnati Division of Criminal Justice

This research was supported by funding from Ohio Office of Criminal Justice Services (2006-JG-EOR-6244). The findings and recommendations expressed within this report are from the authors and do not necessarily represent the official positions of the Office of Criminal Justice Services. Please direct all correspondence regarding this report to James Frank, Ph.D., University of Cincinnati, PO Box 210389, Cincinnati, OH 45221, phone: (513) 556-5832, email: [james.frank@uc.edu](mailto:james.frank@uc.edu)

## **Introduction**

Policing is a very labor intensive field. The nature of police work and organizations is becoming more complex and challenging, especially considering the expectations and demands on officers under community- and problem-oriented policing. Community-oriented policing suggests that policing be decentralized and that officers solve community problems by being both proactive and creative, while involving community members in producing community safety and security (Buerger 1994; Cordner 1995; Mastrofski 1992; Trojanowicz and Bucquerox 1990). With these changes, the quality of police personnel has perhaps become the key element in the effective execution of these police goals (Grant and Grant 1995; Roberg, Kuykendall and Novak 2002). During the past decade agencies have shifted the focus of their selection processes from eliminating the most problematic recruits to identifying those who are most qualified and appropriate for selection (Carter and Radelet 1999).

The general presumption is that the selection of quality personnel translates into effective crime fighting, positive community interaction, and overall improvements in police accountability. In contrast, there are numerous negative consequences for the city and community if qualified officers are not selected, though they are not easy to quantify. In a general sense, the failure to select quality personnel undermines the ability of law enforcement agencies to protect their citizens. In addition, unqualified officers create problems for the agency and community in terms of complaints for excessive force, improper use of coercive activities, and inequitable police practices. Furthermore, the agency endures costs associated with both hiring and training of replacement officers and defending unnecessary and costly lawsuits. Together these factors have deleterious effects on police-community relations. Even more

unfortunate, they often have the most negative consequences in communities that are least able to deal with them.

### **Police Officer Selection and Hiring: Research and Process**

Unfortunately, law enforcement agencies have been unable to successfully develop a system that can identify, during the recruitment process, which individuals will become the most effective officers. Whether this has resulted from the failure of research to identify selection criteria that predict successful police officers or measurement issues related to operationalizing “quality” street-level performance is unknown. In any case, selection systems used by police agencies have not changed substantially over time.

The majority of recent police research that has addressed the selection and performance of officers has focused on either psychological screening tools or Early Warning Systems. Psychological screening tools have been used to select new recruits from pools of applicants using psychological tests that identify unsuitable candidates (Metchik 1999). The emphasis of these efforts has been to “screen out” recruits who possess psychological traits associated with poor police performance. These tests however are not as reliable for “selecting in” officers (Sanders, 2003). Most of the studies examining policing and selection criteria find a relationship between personality traits and negative predictors of police performance and officer success (e.g. problem officers and poor performance) (Sanders, 2003; White, 2008;). At the same time, these tests are not successful at predicting quality performance. For example, Burkhart (1980) found a correlation between low IQ scores and poor performance, though high IQ scores were not correlated with good performance (see also comments by Sanders 2003). Similarly the MMPI which is used as a screening tool by many agencies is primarily a screening out tool that is designed to discover psychological abnormalities (Metchik 1999). Use of these various

psychological screening tools often results in the elimination of the worst candidates without the ability to identify those that remain in the pool as applicants who possess a set of characteristics and traits associated with successful police careers.

Early Warning Systems are intended to use official agency data on problem officers (e.g. complaints for use of force, improper conduct) to identify unprofessional and/or problematic officer behavior, which may warrant some form of intervention early on in the officer's career. The intent of these systems is to identify and then intervene before officer behavior becomes more serious. It is not a tool for predicting future behavior (Walker, Alpert, and Kenney, 2000) based on the characteristics and traits of individual officers. Instead, it is designed to change the behavior of officers already on the force, who due to their prior police conduct appear more likely than others to exhibit problematic behavior in the future.

Law enforcement agencies, including the Cincinnati Police Department, continue to use a "multiple hurdles" approach to selection where the applicant must pass a series of tests. One of these hurdles involves psychological testing. Usually, prior to the psychological testing of candidates, applicants must pass character and medical evaluations and score above a designated threshold on a Civil Service examination. All of these hurdles are also intended to eliminate applicants who do not meet recognized standards for officers (Metchik, 1999). As previously noted, the focus is often on removing applicants who appear to possess undesirable characteristics.

Other areas of the criminal justice system have been much more productive in their efforts to identify potential future success of individuals. Several areas, such as probation and parole, have developed assessment instruments for predicting successful outcomes. Over the years, the statistical sophistication of these prediction devices has greatly improved, making

them more accurate and reliable. In theory, similar techniques should be available to determine which candidates will be the best officers and will enjoy successful police careers. However, in contrast to probation and parole measures of success, successful policing is difficult to operationalize and varies across departments and agency strategies.

### **Measurement of “Quality” Policing**

One problem with measuring effective police behavior centers on the difficult nature of quantifying “quality” while at the same time taking into account the multiple dimensions of police work. This has led to a lack of consensus as to how to measure quality performance on the street (Bartol 1991). Historically, police agencies have used counts of arrests and citations to measure officer productivity. As such, officer evaluations have largely been a “numbers game” (Bayley and Bittner 1984; Fyfe 1999; White 2008), and active officers are viewed as amassing certain levels of productivity.

Unfortunately, these measures, while easy to count, only assess the application of formal authority by officers, and law enforcement has been shown to consume only a limited portion of police work ( Parks, Mastroki, DeJong, and Grey 1999; Smith, Novak, and Frank 2001), even in the era of community policing. In this era, community members now expect more of police officers than just crime control, especially in the era of community policing. An additional limitation to the “numbers game” for measuring quality is that officer arrest and citation activity is likely to be influenced by officer assignment. That is, both variation in shift times and crime levels within the geographical areas officers are assigned to, influence their arrest and citation opportunities (Bayley and Bittner 1984; Fyfe 1999; Muir 1977; Skolnick and Fyfe 1993). Finally, the number of arrests and citations may be influenced by officer choices to produce quantifiable numbers versus selecting the appropriate intervention.

There are also qualitative dimensions to “good officers.” As Sanders (2003) comments, it is much easier to recognize and measure bad performance than good performance. While there is a lack of consensus as to what is good, it is commonly recognized that nonfeasance, excessive use of force, discourteous and aggressive actions, as well as other behaviors which undoubtedly cross the line into improper police conduct represent traits of bad officers. However, quality police action may depend on characteristics of the specific encounter and the people involved. For instance, in an intervention with a citizen is an arrest always better than a warning? Does the selection of one action produce consistently better final outcomes than another?

Finally, measuring success is further complicated by issues surrounding data availability. Many agencies rely on annual personnel assessments made by supervisors that are designed to account for street-level enforcement activity, compliance with agency rules, and general work history. Two criticisms have been regularly voiced concerning annual evaluation processes. First, questions have been raised over whether the evaluation instruments measure the tasks officers regularly perform during their typical work shift. For instance, are assessments flexible enough to adapt to the changing demands placed on officers under community policing? Second, there have always been questions about whether ratings reflect supervisor perceptions or actual police performance (Doerner and Hunter 2006).

### **The Present Study**

The challenge of selecting quality police personnel has been a critical problem for law enforcement agencies in Ohio, especially Cincinnati. A report comparing Ohio's 1981 and 1996 Task Analysis data indicated that almost a third fewer officers (96% compared to 66%) claim that they are assigned solely to patrol duties (Travis and Sanders 1997). Further, the Cincinnati Police Department, in particular, has made great efforts in the last decade to create a more

community-oriented police department. As a result, it is imperative that new recruits are both highly qualified and dynamic. As policing on the street becomes more complex, there is an even more urgent need to select qualified personnel.

Several years ago, Ohio police administrators identified their most pressing information need as a response to the question, "Are there personal attributes that can be designed into the recruit applicant procedures that predict effective future officers?" In a series of regional research forums conducted by the Law Enforcement Foundation, in cooperation with the Ohio Association of Chiefs of Police, Ohio police chiefs also identified the third highest ranked research question as, "What are the predictors of successful police officer performance and how do we select for them?" Unfortunately, there is a lack of evidence illuminating which criteria are the best predictors of quality personnel.

The present study seeks to employ techniques similar to those used in other areas of the criminal justice system to examine information that police agencies have at their disposal during the selection and academy stages in order to predict which recruits will successfully adjust to the duties of policing. To that end, the focus will center on three primary research questions: (1) Do personal qualifications predict performance in the training academy? (2) Does performance in the training academy predict success as an officer? (3) What are the most appropriate and effective measures of officer success? While some limited research concerning the relationship between selection criteria and academy performance exists (White 2008), the relationship between academy performance and street-level quality is largely untested. This may result from the difficulties associated with defining and operationalizing street-level success and quality performance. The findings of this examination will enable police administrators to better develop

reliable selection systems that can be used to screen for higher quality candidates rather than the more traditional methods of screening for negative attributes.

## **Methods**

Utilizing secondary data analysis, the main focus for this study is the relationship between individual officer characteristics and various dimensions of quality officer performance. Data were collected from the Cincinnati Police Department from June 2007 through January 2008. The Cincinnati Police Department is a medium-sized department with approximately 1,050 sworn officers across 5 districts, serving about 330,000 citizens (Uniform Crime Report 2007; U.S. Census Bureau 2000). Data were collected for all officers who entered the police department's training academy program from 1996 to 2006 and eventually served with the agency. Individuals who resigned from the academy, failed to fulfill the requirements of the training academy, and/or ultimately were not hired for police work were removed from the sample. Likewise, those individuals who were listed on the academy roster, but had no performance information recorded were treated as missing cases. That is, persons admitted to the academy who did not actually enter, complete and/or join the police force were not included in the sample of officers. As a result of these conditions, approximately 30 officers were removed from the initial sample. The final sample consisted of 486 Cincinnati police officers.

### **Data Collection**

When first contacted by the Cincinnati Police Department, we were provided limited personal data on officers who entered their police academy during the years 2001-2006. These data were supplemented during the data collection efforts described below with additional personal, academy, and performance data. In addition, data were also collected on officers who entered the academy and joined the police force during the previous five years (1996-2001).



Given the type and amount of information necessary to perform the desired analysis, data had to be collected from several separate offices within the police department—the training academy, the department’s personnel office, and the department’s internal investigation unit. In addition (as is discussed below), efforts were made to collect information from the Crime Analysis Unit. At each location, data were collected directly by the research team members or were provided in raw form by support staff within the police agency office. In most cases, relevant records were transmitted from hard copies provided onsite to the research team’s laptop computers. One exception was the information obtained from the internal affairs office, which was provided on a compact disc (for security reasons we were not given access to internal affairs files). Data from the Crime Analysis Unit was also provided on a compact disc .

#### *Data from the Police Academy*

The majority of the research data were obtained from the training academy. These records included both officer demographic information and academy performance data. At the training academy, annual academy class information is kept in large binders, stored in the file room. Research team members examined and recorded information from these binders. On average, there were two academy classes per year. As such, data were collected from 18 separate binders. The demographic information collected included both typical personal data pertaining to officer characteristics such as gender, race, age, educational level and foreign language skills, as well as more specific officer-level data such as prior military and law enforcement experience. The academy performance data included both academic test scores (quizzes over substantive areas, final test score, spelling grades, notebook organization, etc) as well as physical performance scores (number of sit-ups, number of push-ups and mile and a half run times within specified time periods). A hard copy of Civil Service test scores was also

provided by representatives of the police academy after they were retrieved from files located at a separate city office building. Unfortunately, we were only able to collect Civil Service scores for a portion of the officers (N=341).

#### *Data from the Personnel Office*

Information describing the performance of officers during active service was obtained from the personnel office. As with the training academy information, the data collected from the personnel office were collected onsite by members of the research team. Data were coded and recorded directly from hard copy personnel files. The data retrieved from the personnel office included annual performance evaluation scores, information pertaining to officer auto accidents, as well as promotion and officer commendation information. More specifically, the promotion information was restricted to whether the officer secured a promotion during the study period or not. Data were also collected on two types of commendations--official commendations issued by police department superiors and citizen-based commendations due to community member reports concerning officer behavior.

#### *Data from Internal Investigations*

Given the relatively private nature of much of the internal investigations information, the research staff was not granted access to the Internal Investigations' files. Instead, a list containing the names and badge numbers of the officers in the sample was provided to a member of the Internal Investigations' support staff. The staff member then provided the investigators with an electronic file listing the types of complaints and the outcomes of complaint investigation processes for all officers in the sample. The research team transferred and coded the data from the electronic file provided into the study database. Complaint data included information on the nature of the complaint (e.g. a complaint of excessive use of force, racial

profiling, failure to appear in court, etc.), as well as disposition of the complaint (e.g., the officer was exonerated, the complaint was determined to be unfounded, the complaint was sustained, etc.).

#### *Data from Crime Analysis Unit*

Data on officer arrest activity was secured from the Crime Analysis Unit. Specifically, we requested information on all arrests for Part 1 and Part 2 crimes during the study's ten year period. This information was provided to us on a diskette and was then sorted by badge number and officer name. The intent was to utilize this information as a measure of street-level officer activity, while also accounting for officer assignment and years on the force. Unfortunately, due to missing information, we were not able to use this data in our analyses. Issues surrounding incorporation of this data are discussed again in the section on dependent variables and are revisited in the discussion section of this report.

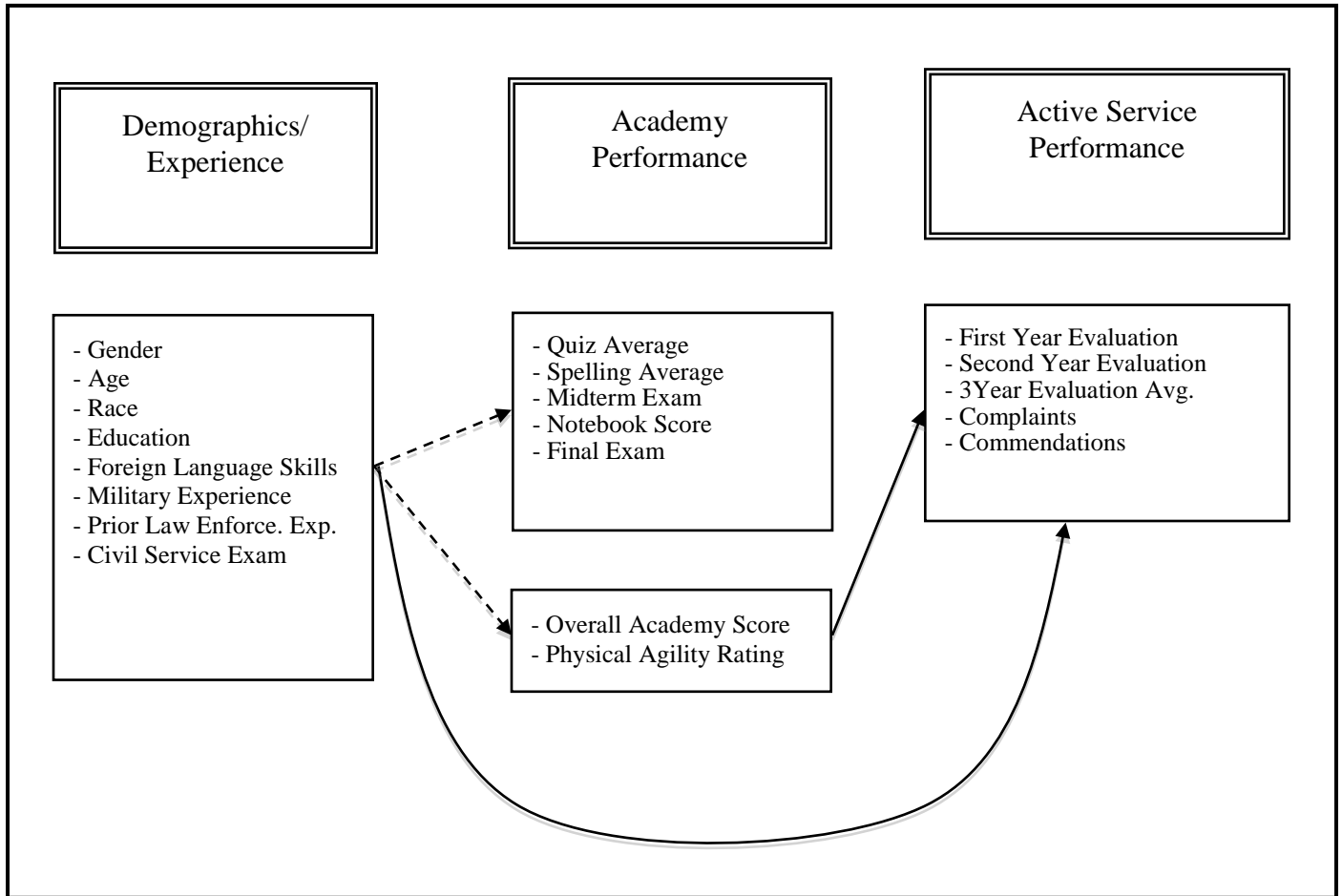
In summary, all of the collected information, except the arrest data, was combined into a single officer-level database. The data file contained all personal information available at the time officers applied to the police agency, data generated during the academy, and performance measures once a member of the police force. Unfortunately, psychological testing scores were not available. The collected measures were used to develop models that were used to explore relationships between officer demographic variables and both academy performance, as well as performance while on the police force.

#### **Analytic Strategy**

For the current examination, a two stage analysis technique was used. First, we examined the relationship between the officer demographic/experience variables and the academy performance variables. Next, we examined the relationship between the officer

demographic/experience variables, and the two major academy performance variables—overall academy score and physical agility rating—and the active service performance variables. As performance during active service is one of the key measures of a quality officer, the active service performance variables represented the measures of success as an officer.

**Figure 1: Path Model for Data Analysis.**



That is, for this investigation, the demographics/experience variables, as well as the academy performance variables, were utilized as predictors of future officer success. This technique was adopted in order to determine if both direct and indirect correlations exist.

The basic path model adopted can be seen in Figure 1. The dashed lines represent the stage one analysis, and the solid lines represent the stage 2 analysis. For both stages of the

analysis, multivariate regression was performed in order to examine all possible correlations between the independent and dependent variables. Before estimating equations to predict officer success, a number of exploratory analyses were conducted that primarily entailed computing correlation coefficients. This was necessary because the assessment of relationships between collected hiring criteria and officer quality is one of first impression for the police agency. So, it was important that relationships be examined prior to developing a parsimonious model predicting officer success.

### **Dependent Variables**

The dependent variables examined in this study include both academy performance measures and active service performance measures. As previously noted, scores generated during the police academy are the dependent variables for the first stage of the analysis, while department-generated measures serve as the dependent variables during the second stage of the analyses (see Figure 1).

For the first stage of the analysis, academy performance measures include the following training academy scores: mean score for quizzes, mean score for spelling exams, midterm exam score, notebook score, final exam score, and an overall final academy score. These scores were collected from the academy class binders, and each of these measures was coded as a metric variable.

An additional academy score was calculated using physical agility information provided in the academy class binders. The physical agility rating includes measures of timed running, push-ups, and sit-ups. The individual outcomes for each physical agility test were ranked, recoded, and combined to create an overall physical agility rating, since it was presumed that overall physical conditioning was more important than performance of just one of the three types

of physical activity. Specifically, for each activity (sit-ups, push-ups, and run) a mean officer performance score was first calculated and assigned a score of 4. For each standard deviation above the mean an additional point was added and the maximum score became a 7. A point was deducted for each standard deviation below the mean. For each physical activity there was a possible score that ranged from 1 to 7, and the total physical agility rating score of 3-21 for all three measures. The scales and descriptive measures for the academy performance measures can be seen in Table 1.

The creation of the dependent variables for the second stage of the analysis was much more difficult. One of the most critical issues in the selection and hiring literature involves questions about how to measure and operationalize officer success and quality. Since this is a question of first impression, we explored a number of measures of success in an effort to create a measure that encompassed a variety of dimensions of police work. In creating the active service performance measures, variables were used that allow for examination of both department and community perceptions of officer activity. As many law enforcement agencies, including the Cincinnati Police Department in particular, are pushing towards a community-oriented style of policing, it is important to include the perceptions held by those in the community regarding the police. Full descriptive statistics for all of the dependent variable measures can be seen in Table 1.

Included in the performance measures that independently describe the department's perception of officer performance are the annual evaluation scores. As previously stated, annual evaluation scores were collected for each officer in the sample. Use of these scores in a meaningful manner created several challenges. First, for officers in the 1996 academy class, we had at least ten evaluation scores, while for the 2006 class of officers only one or two scores

were available. Second, initial examinations of the score frequencies highlighted the fact that scores tended to increase for most officers over their tenure on the force. For instance, mean officer evaluation scores increased almost two points when the first and second year evaluations are compared (11.861 versus 13.730), a trend that remained until officers tended to level off at a higher score after several years (see Table 1). To retain a sufficient number of officers in our sample and to make comparisons across officer evaluations, it was necessary to limit the number of annual evaluation scores used in the analyses. Specifically, we decided to use only scores from the officer's first and second years of service, as well as the average evaluation score for the officers' first three years of service. That is, our analysis using evaluation scores permits us to examine the impact of officer-level characteristics and academy performance scores on their department performance evaluations for their first several years of active service.

In order to examine the community's perception of officer performance, a variable was also created which measured the number of complaints filed against individual officers. While some of these complaints could have been filed by members of the department, the overwhelming majority of them were filed by citizens within the community. Complaints vary widely in terms of the nature of the complaint against the officer. The typology utilized in this examination includes complaints for: misuse of force, racial profiling, missing court appearances, unlawful activity (such as driving under the influence, domestic violence, etc.), legal issues (such as illegal searches, etc.), sexual misconduct, and an "other" category (which consists of complaints of unprofessional behavior, cursing, etc.).

It was ultimately decided to use the total number of complaints against the officer, rather than measures of individual types of complaints. This decision was made primarily because most of the categories of complaints did not have sufficient numbers to allow for separate

statistical analysis. For example, the most common type of complaint was based on improper use of force. Though the most common, they were rare. The mean total number of use of force complaints for officers was 1.2, and 56 percent of the officers were not the subject of such a complaint. Almost thirty percent (29.3%) of the officers had one or two complaints lodged against them. In contrast, the other types of complaints were much less likely to occur as evidenced by their mean frequency scores (racial discrimination = .07; legal complaints = .37; missing court complaints = .43; law violations = .26; sexual misconduct = .06). Second, in an attempt to eliminate any possible time period or history effects, the number of complaints filed against officers needed to account for years on the force. To accomplish this, the number of complaints was standardized by the number of years of service as an officer. Thus, the complaints measure is the total number of all complaints standardized by years on the force.

Finally, a measure that combines both department and community perception of performance was also included—commendations. Commendations are received for a variety of reasons and may be initiated by citizens or department personnel. Citizens may send letters to the police department in response to a positive encounter with an officer (i.e., the officer helped me unlock my car), which may ultimately result in a commendation. If an officer exceeds expectations in the course of his/her duties (i.e., the officer voluntarily led a neighborhood crime prevention meeting), commendations may also be generated by department supervisors. Finally, commendations can be the result of bravery in the line of duty or some other meritorious behavior. Citizen-based commendations were much more common than official department commendations. For example, 52.5 percent of the sample did not receive an official commendation, while only 21.8 percent of the officers were without a single citizen-based commendation.



As with complaints, the number of commendations for individual officers was standardized by the officer's total number of years of service with the police agency. This decision was made because each year on the force represents an additional opportunity for officers to receive commendations. Thus, to account for differences in opportunity there was a need to standardize this measure. In addition, specific years seemed to produce an unusual number of commendations across the entire department.

An attempt was also made to measure officer activity on the street by examining the number of arrests by each officer in the sample. When the data was cleaned and sorted by officer badge number, we encountered a substantial problem involving missing data. Specifically, two critical problems emerged. First, for each year in our dataset between 25 and 30 percent of the arrests were missing badge numbers. Second, when arrest entries that contained badge numbers were examined there were often multiple names (1 to 5) associated with the badge number in a given year. Our intent was to match the arrest information with the officer-level data in our data set. Since a considerable portion of the arrest related information could not be matched (almost one-third of the data in each study year), we did not have confidence in our ability to create a meaningful measure of this activity using the provided data.

### **Independent Variables**

The independent variables examined in this study consisted of both general demographic information and more specific police-related information. The individual officer characteristics examined included: gender, age at time of recruitment, race, college education, foreign language skills, military experience, prior law enforcement experience, and Civil Service exam scores. Gender, race, college education, foreign language skills, and military experience were each coded as dichotomous yes/no measures. With college education, any college experience—not

just obtaining a college degree—was counted as college education. Age at time of recruitment, prior law enforcement experience, and Civil Service exam scores were each coded as metric variables, representing the actual number of years or raw scores respectively. Unfortunately, there were no available data describing officer personality characteristics. As a result, our examination is limited to demographic, experience, and performance data. A more detailed description of these variables can be seen in Table 1.

## **Results**

This section reports the results of the data analyses using the officer-level hiring and performance data. Initially, correlation coefficients were computed to explore relationships among the variables and to inform the multivariate analysis. Thereafter, the relationships were explored further by developing models used to predict academy and performance success.

### **Correlation Analysis**

This analysis follows the strategy of first examining academy outcomes and then active service measures. More specifically, we first present the bivariate correlations between officer characteristics and academy scores. We follow this with the bivariate correlations between all variables and several service measures. Correlation coefficients are a summary measure of the covariation between two variables. The coefficient measures the association between two variables, without any other “control” variables, and indicates the existence and strength of the relationship.

**Table 1: Variables, Scales, and Descriptive Statistics**

<b>Variables</b>	<b>Scale</b>	<b>Mean</b>	<b>S.D.</b>	<b>Range</b>	<b>N</b>
<b>Independent Variables</b>					
Gender	(0 = Male, 1 = Female)	0.210	0.406	0 – 1	486
Age	(Age at Recruitment)	29.210	5.555	21 – 55	486
Race	(0 = White, 1 = Nonwhite)	0.370	0.483	0 – 1	486
Education	(0 = HS diploma, 1 = College Experience)	0.854	0.354	0 – 1	486
Foreign Language Skills	(0 = No, 1 = Yes)	0.080	0.278	0 – 1	486
Military Experience	(0 = No, 1 = Yes)	0.310	0.464	0 – 1	486
Prior Law Enforcement Exp.	(Number of Years of Prior Experience)	0.473	1.261	0 – 8	486
Civil Service Exam	(Civil Service Exam Score)	87.539	6.719	69.00 – 101.47	341
<b>Dependent Variables</b>					
Quiz Average Score	(0 – 100)	86.538	6.054	61.60 – 99.20	458
Spelling Exam Score	(0 – 101)	93.703	5.809	60.00 – 100.67	407
Midterm Score	(0 – 100)	85.073	6.975	60.70 – 98.67	432
Notebook Score	(0 – 100)	87.120	11.069	38.09 – 100.00	290
Final Exam Score	(0 – 100)	83.978	6.262	61.38 – 98.66	376
Overall Academy Score	(0 – 100)	86.439	5.710	70.00 – 97.21	420
Physical Agility Rating	(0 – 21)	12.265	3.315	2.00 – 21.00	358
First Evaluation Score	(0 – 25)	11.861	1.334	8 – 20	472
Second Evaluation Score	(0 – 25)	13.730	1.459	10 – 21	440
3 Year Evaluation Average	(0 – 25)	13.637	1.230	9.5 – 21	396
Complaints	(Total Number of Complaints)	0.229	0.706	0 – 14.00	486
Commendations	(Total Number of Commendations)	0.555	0.725	0 – 5.82	485

As can be seen in Table 2, two variables have significant relationships with each of the academy measures. That is, Civil Service test scores have a positive relationship with each of the six academy scores, so that higher Civil Service scores are associated with higher academy scores. In contrast, officer race has a negative relationship with each of the scores, indicating that being non-white was associated with lower scores. Education level had a significant and positive association with five of the six scores. The physical agility measure was positively associated with three of the academy measures, while officer gender was negatively related to three academy outcomes. Finally, age and prior law enforcement experience were not significant with any of the academy outcome measures. Overall, the coefficients indicate that most of the observed relationships are weak to modest at best.

**Table 2: Bivariate Correlations Between Officer Characteristics and Academy Outcome Measures**

<b>Variable</b>	<b>Quiz</b>	<b>Spelling</b>	<b>Midterm</b>	<b>Notebook</b>	<b>Final</b>	<b>Overall</b>
Gender	NS	-.134*	-.106**	NS	-.102**	NS
Age	NS	NS	NS	NS	NS	NS
Race	-.509*	-.332*	-.453*	-.323*	-.440*	-.526*
Education	.267*	.257*	.264*	NS	.308*	.321*
Foreign Lang	NS	NS	.122**	NS	.144*	NS
Military Exp	.093**	NS	NS	NS	NS	NS
Prior Law Enforce Exp	NS	NS	NS	NS	NS	NS
Civil Service	.510*	.394*	.530*	.296*	.621*	.617*
Physical Agility Rating	NS	.154**	.122**	NS	NS	.113**

\*  $p \geq 0.01$ ; \*\*  $p \geq 0.001$

Table 3 presents the bivariate correlations between the officer and academy measures and active service outcomes. Several findings are important to note. First, only three of the variables are significantly related to the first year evaluation scores. Prior law enforcement experience and higher physical agility scores are positively associated with officer evaluations in their first year. Officer gender (being female) was negatively associated with first year evaluation scores. Second, in contrast, many more variables were positively related to the second annual evaluation scores, and all but two of the relationships (officer gender and race) were positive. Third, both officer gender and physical agility ratings were significantly related to all three annual evaluation scores.

**Table 3: Bivariate Correlations Between Officer Characteristics and Academy Outcome Measures and Evaluation Scores**

<b>Variables</b>	<b>Year 1 Evaluation</b>	<b>Year 2 Evaluation</b>	<b>3 Year Evaluation Avg</b>
Gender	-.151*	-.169*	-.219*
Age	NS	NS	.107**
Race	NS	-.139*	-.110
Education	NS	NS	NS
Foreign Language	NS	NS	NS
Military Experience	NS	.133*	.119*
Prior Law Enforcement Exp	.110**	NS	.179*
Civil Service Exam	NS	.190*	NS
Physical Agility Rating	.115**	.146*	.188*
Quiz Average	NS	.208*	.181*
Spelling Average	NS	NS	NS
Midterm Exam	NS	.153*	NS
Notebook Score	NS	.172*	.233*
Final Exam	NS	.207*	NS
Overall Academy Score	NS	.198*	.116**

\*  $p \geq 0.01$ ; \*\*  $p \geq 0.001$

Table 4 and Table 5 present bivariate correlations between the same officer variables and alternative measures of officer performance in the field. Table 4 reports the correlations with both commendation measures and complaints filed against officers. Only a limited number of the computed relationships are statistically significant. Of interest are the findings that prior law enforcement and military experience were positively related to more total complaints, though when complaints are standardized by years on the force the relationships disappear. Civil Service test scores were negatively related to complaints suggesting that officers with higher scores were less likely to have large numbers of complaints filed against them. Only two variables exerted relationships across both complaint measures (gender with women having fewer complaints and higher physical agility scores related to more complaints). Table 5 reports the bivariate correlations with number of auto accidents and use of force complaints. Again, only a few of the computed relationships are significant.

Several summary findings suggest that the predicting active service performance may be more difficult than academy performance. First, only a limited number of officer measures were consistently significant with the annual evaluation scores. Second, only a limited number of predictor variables were significantly related to the alternative service outcomes. Third, coefficients for relationships that were statistically significant indicate that the associations were modest at best, while most would be considered weak.

### **Multivariate Models**

Multivariate regression analysis was used in order to assess the independent effects of included variables. As can be seen in Table 6, several key findings stand out in the first stage of the analysis. First, there is a consistent relationship between officer race and academy performance. This relationship is significant and negative for every measure of academy

performance. Essentially, nonwhite officers did not perform as well in the academic aspects of the training academy. Potential reasons for this finding will be discussed in more detail later. There is also a significant relationship between prior law enforcement experience and several of the academy performance measures. The relationships between prior experience and quiz average, final exam score, and overall academy score are both significant and positive, indicating that individuals with prior law enforcement experience perform better on the training

**Table 4: Bivariate Correlations Between Officer Characteristics and Academy Outcome Measures and Commendations and Complaints**

Variables	Officer Commendations	Officer Commendations Std	Official Commendations	Official Commendations Std	Complaints	Complaints Std
Gender	NS	NS	-.139*	-.091**	-.198**	-.209**
Age	.205**	NS	NS	NS	.093*	NS
Race	NS	NS	-.114**	-.097**	NS	NS
Education	NS	NS	NS	NS	NS	NS
Foreign Language Skills	NS	NS	NS	.119*	NS	NS
Military Experience	.120*	NS	.107**	NS	.184**	NS
Prior Law Enforcement Exp	.165*	NS	.124*	NS	.161**	NS
Civil Service Exam	NS	NS	NS	NS	-.141**	NS
Physical Agility Rating	NS	NS	NS	.104**	.134*	.129*
Quiz Average	.126	NS	.182*	NS	NS	NS
Spelling Average	NS	NS	NS	NS	-.128**	NS
Midterm Exam	NS	NS	.148*	.105**	NS	NS
Notebook Score	NS	NS	NS	NS	NS	NS
Final Exam	NS	NS	NS	NS	NS	NS
Overall Academy Score	NS	NS	.118	NS	NS	NS

\*  $p \geq 0.01$ ; \*\*  $p \geq 0.001$



**Table 5: Bivariate Correlations Between Officer Characteristics, Auto Accidents, and Use of Force Complaints**

<b>Variables</b>	<b>Auto Accidents</b>	<b>Auto Accidents Std</b>	<b>Use of Force Complaints</b>
Gender	NS	NS	-.178
Age	.157**	NS	NS
Race	NS	NS	NS
Education	NS	NS	NS
Foreign Language Skills	NS	NS	NS
Military Experience	.209**	NS	.160**
Prior Law Enforcement Exp	.089**	NS	.133*
Civil Service Exam	NS	NS	NS
Physical Agility Rating	NS	NS	.127**
Quiz Average	.131*	.100**	NS
Spelling Average	-.103**	NS	NS
Midterm Exam	NS	NS	NS
Notebook Score	NS	NS	.128**
Final Exam	NS	NS	NS
Overall Academy Score	NS	NS	NS

\*  $p \geq 0.01$ ; \*\*  $p \geq 0.001$

training academy measures . Finally, and possibly most importantly, a positive and very significant relationship was found between Civil Service exam scores and every measure of academy performance. This indicates that individuals who score higher on the Civil Service exam perform better in the training academy. This conclusion supports the purpose of the Civil Service exam—to determine which individuals have the abilities for potential future success. These findings provide some support for the first stage of the analysis, as presented in Figure 1, especially since the R squares suggest some success in explaining model variation. Certain individual demographic/experience measures (especially Civil Service scores) can be used to predict, with some accuracy, potential success in the Cincinnati police training

**Table 6: OLS Regression Analysis between Individual Demographic/Experience Characteristics and Academy Performance Measures**

Variables	Quiz Average	Spelling Average	Midterm Exam	Notebook Score	Final Exam	Overall Academy Score
Gender	0.421 (0.625)	1.490 (0.784)	-0.091 (0.792)	4.822 (1.805)**	-0.320 (0.723)	0.546 (0.573)
Age	-0.014 (0.044)	-0.012 (0.056)	-0.123 (0.056)***	-0.036 (0.122)	0.047 (0.049)	-0.042 (0.040)
Race	-3.567 (0.612)*	-2.114 (0.793)**	-2.811 (0.777)*	-5.172 (1.941)**	-2.345 (0.678)*	-3.064 (0.556)*
Education	0.422 (0.773)	1.664 (0.983)	1.218 (0.975)	-0.700 (2.462)	1.527 (0.878)	1.014 (0.717)
Foreign Language Skills	-1.323 (0.935)	-0.384 (1.218)	1.132 (1.179)	-0.540 (2.719)	0.341 (1.077)	-0.621 (0.853)
Military Experience	0.246 (0.556)	-0.523 (0.753)	0.158 (0.702)	-0.119 (1.887)	-1.593 (0.615)**	-0.607 (0.503)
Prior Law Enforcement Exp	0.753 (0.234)*	-0.922 (0.462)***	0.448 (0.298)	2.666 (2.345)	0.567 (0.257)***	0.547 (0.222)***
Civil Service Exam	0.291 (0.044)*	0.259 (0.056)*	0.420 (0.056)*	0.394 (0.144)**	0.498 (0.051)*	0.403 (0.041)*
R <sup>2</sup>	0.345	0.192	0.332	0.126	0.427	0.455

\*  $p \geq 0.05$ ; \*\*  $p \geq 0.01$ ; \*\*\*  $p \geq 0.001$

academy, though most of the variables in the model were not significant predictors of the academy success outcomes.

For stage 2 of the analysis, the demographic/experience measures and the major academy performance measures—overall academy score and physical agility rating—were examined in relation to active service performance measures. Again, multivariate regression analysis was used to assess the independent effect of variables in the model while controlling for other factors. As can be seen in Table 7, only a few correlations prove to be significant in this stage of the analysis. First, in examining the relationships between the individual demographic/experience measures and the active service

**Table 7: OLS Regression Analysis between Individual Demographic/Experience Variables, Academy Performance Variables, and Active Service Performance Variables**

Variables	Year 1 Evaluation		Year 2 Evaluation		3 Year Evaluation Average		Complaints		Commendations	
Gender	-0.376 (0.193)***	-0.467 (0.235)***	-0.613 (0.221)**	-0.548 (0.277)***	-0.367 (0.209)	-0.389 (0.254)	-0.414 (0.094)*	-0.321 (0.122)**	0.066 (0.122)	-0.073 (0.144)
Age	-0.032 (0.013)***	-0.037 (0.014)***	-0.028 (0.015)	-0.019 (0.017)	-0.036 (0.014)***	-0.037 (0.016)***	-0.023 (0.007)*	-0.024 (0.008)**	0.011 (0.008)	0.009 (0.009)
Race	-0.040 (0.185)	0.296 (0.217)	-0.134 (0.205)	0.009 (0.254)	-0.044 (0.185)	0.128 (0.229)	-0.019 (0.090)	-0.101 (0.113)	0.054 (0.117)	0.136 (0.133)
Education	0.149 (0.230)	-0.015 (0.267)	-0.135 (0.254)	-0.119 (0.310)	0.005 (0.232)	-0.015 (0.286)	-0.118 (0.112)	-0.150 (0.138)	0.104 (0.145)	0.141 (0.163)
Foreign Language Skills	-0.695 (0.298)***	-0.695 (0.334)***	-0.132 (0.340)	-0.283 (0.388)	-0.388 (0.317)	-0.483 (0.362)	0.119 (0.141)	-0.081 (0.172)	-0.198 (0.182)	-0.168 (0.203)
Military Experience	0.286 (0.167)	0.132 (0.200)	0.176 (0.186)	0.027 (0.235)	0.218 (0.171)	0.158 (0.215)	0.127 (0.082)	0.250 (0.105)***	-0.131 (0.106)	-0.053 (0.123)
Prior Law Enforcement Exp	0.124 (0.071)	-0.007 (0.096)	0.078 (0.076)	0.055 (0.111)	0.060 (0.066)	-0.028 (0.095)	0.013 (0.035)	0.000 (0.050)	0.050 (0.045)	0.097 (0.059)
Civil Service Exam	0.016 (0.014)	0.004 (0.019)	0.036 (0.015)***	0.011 (0.022)	0.032 (0.014)***	-0.007 (0.020)	-0.010 (0.007)	-0.011 (0.010)	-0.002 (0.009)	0.000 (0.012)
Overall Academy Score		0.043 (0.022)***		0.027 (0.026)		0.061 (0.025)***		-0.004 (0.012)		-0.003 (0.014)
Physical Agility Rating		-0.008 (0.032)		0.033 (0.037)		-0.012 (0.033)		0.021 (0.017)		0.003 (0.020)
R <sup>2</sup>	0.055	0.053	0.061	0.029	0.061	0.048	0.083	0.098	0.000	0.000

\* p ≥ 0.05; \*\* p ≥ 0.01; \*\*\* p ≥ 0.001

performance measures, three specific relationships are significant. Gender is shown to be significantly and negatively associated with both the evaluation score variables and the variable measuring the number of complaints. These findings indicate that female officers generally receive lower evaluation scores; however, they also receive fewer complaints. A similar relationship is seen with the variable measuring age at time of recruitment. As can be seen in Table 7, officers who were older at the time of recruitment generally received lower evaluation scores. Also, as with female officers, they tend to have fewer complaints filed against them.

Two additional variables are also worthy of mention. There is some evidence that Civil Service scores are positively and significantly associated with evaluation scores. However, the relationship is no longer significant once the officer's overall academy performance score and physical agility rating are included in the models. In these situations the Civil Service exam score does not retain its significance. Second, when the correlations between the academy performance variables and the active service performance variables are examined, the only significant finding is the positive relationship between the overall academy score and two of the evaluation measures. This finding indicates that officers who performed well during the academic portion of the training academy generally receive higher evaluation scores.

Taken as a whole, these findings provide limited support for the path model previously described. First, the reported R squares suggest that the models are only explaining a limited proportion of the variance in the models. Second, most of the individual demographic/experience variables are not significantly related to the active service measures. The exceptions as noted are officer gender and age at time of recruitment. Third, where relationships are observed, many are not consistent across the officer performance outcomes.

Fourth, of all the variables examined in the second stage of the analysis, only commendations were not significantly related to any of the predictor variables. The limited variation in the number of commendations given could be responsible for lack of significance and indicates that commendations may not be an effective measure of officer success.

## **Conclusion and Discussion**

The purpose of this study was to answer three primary research questions; each of these will be answered and discussed below. The first research question: Do personal qualifications predict performance in the training academy? There is some support for the assertion that personal characteristics or qualifications are related to success in the police training academy. Two variables were consistently related to the academy success measures—race and Civil Service exam score. One possible explanation for the effect of the race variable may be that the model is misspecified, and that the variable is instead masking an effect of socioeconomic status, which unfortunately was not considered in the analysis due to a lack of information tapping that dimension. Another possibility is simply that the effect is an artifact unique to this data or this model since the effect disappears in the second model. The predictive power of the Civil Service exam variable makes intuitive sense since this is one of the first screens designed to aid in the hiring of quality police officers. A third variable measuring prior law enforcement experience was moderately successful at predicting success in the academy. This relationship is expected considering those with prior law enforcement experience also are likely to have had prior police academy experience as well.

The second research question: Does performance in the training academy predict success as an officer? There is mixed support for the claim that success in the training academy may equate to success as a police officer. The academy score variable, which is meant to tap overall

success in the academy, is significantly related to the officers' first evaluation and their three year evaluation average. The physical agility score variable was not significantly related to any measure of active service success. Interestingly, neither of these academy variables was significantly related to complaints or commendations. This suggests that perhaps these variables are not the most desirable measures of officer success. The only variables which were related to complaints and commendations were gender, age, and military experience. Female officers received fewer complaints against them than did their male counterparts, while younger officers and those with military experience received more complaints against them. A possible explanation for this could be that female officers are inherently more adept at resolving conflicts, thereby avoiding the possibility of a citizen complaint. Another possibility may be that assignments and patrol areas vary by gender or age; unfortunately, this information was not measured in this study. The receipt of commendations by officers was not significantly related to any variable in the model.

Other interesting findings involve those variables which are often considered good predictors of officer success. Higher education is often thought of as a desirable quality for incoming police recruits to possess. However, education did not prove to be related to any of the measures of academy or on the job success used in these analyses. Similarly, foreign language skills were equally unable to predict academy or officer job success. The only significant effects of military experience were related to the final exam variable and the complaints variable. Prior law enforcement experience adequately predicted academy success, but had no effect in the officer performance model. As discussed, Civil Service scores were a good predictor of academy success; they were also significantly related to officers' second evaluation and three

year evaluation average. This suggests that one of the best criteria for police departments to base hiring decisions on may be high Civil Service scores.

The third research question: What are the most appropriate and effective measures of officer success? This is likely the most difficult question due to the ambiguity involved in police work. For example, number of complaints would appear to be a valid measure of officer quality, as an officer with many complaints would not likely be viewed as a quality officer. However, complaints may actually indicate that the officer is active and willing to interact with citizens, factors which increase the likelihood of a complaint being lodged against an officer. In contrast, few complaints may indicate that the officer is unwilling to intervene.

We attempted to differentiate between complaints that were unfounded versus those sustained by the department. The assumption was that officers with sustained complaints are likely more problematic than officers who have fewer sustained complaints and also those officers with unfounded complaints. Unfortunately, no significant relationships were discovered using these alternative measures.

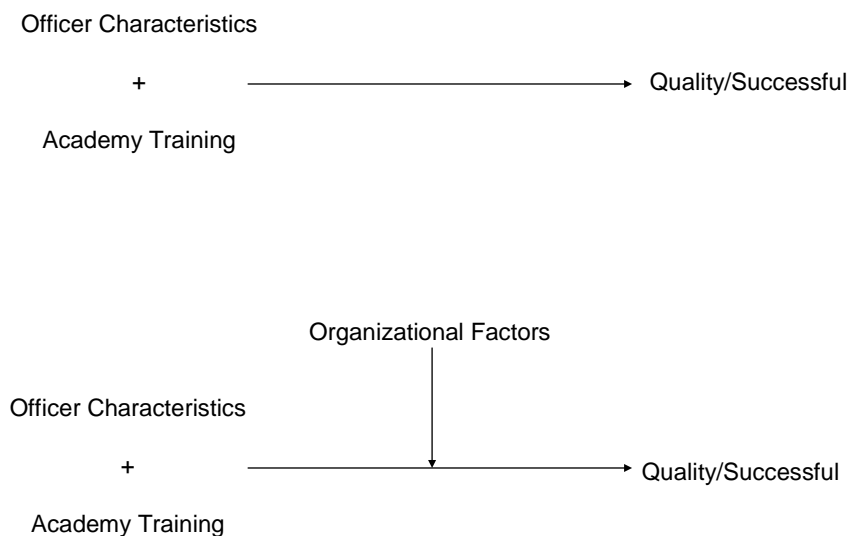
Measuring quality by counting the number of arrests and citations by officers also has problems. As has been suggested by others, the number of officer arrests is influenced by shift assignment and neighborhood assignment. Furthermore, arrest counts are likely influenced by discretionary choices to engage in arrest behavior versus using other strategies when interacting with citizens. Even with these potential problems, citation and arrest activity should be considered only one measure of quality street behavior. If available, other measures of street activity (citizen meetings, dispute resolution, etc.) should be used to supplement counts of more formal actions, especially in the community policing era. We attempted to create a measure of quality that accounted for the various dimensions of police work. Specifically, an attempt was

made to combine service outcomes through a point system. Unfortunately, the measure was not related to most of our officer or academy measures.

Several additional considerations are worthy of mention concerning the selection process and the hiring of quality officers. First, there is a need to reform the process so that selection criteria are related to success, if possible. This may require developing hiring criteria that are empirically related to the tasks that officers actually perform on the street. The purpose would be to create a validated job-related selection process.

Second, the department may need to determine whether to continue the selection process as it is now conducted. The general presumption as seen in Figure 2 is that officer characteristics and academy behavior will predict success. Unfortunately, our findings suggest that most of the information collected during the application process and information generated during the academy is not related to the service outcome measures used in our study. This may be due to organizational factors that intervene after completion of the academy; so that organizational factors mediate the effects of the officer-level factors (see Figure 2). If so, then it may be important to identify and examine those organizational factors that intervene and

**Figure 2: Predicting Officer Success**





influence officer success. Alternatively, it may be that information associated with the hiring of quality officers is not collected during the present hiring process. If this is correct then efforts should be made to determine whether this information can be identified and collected. Finally, it is possible that the necessary information is not available and/or easily collected. As one lieutenant advised us, if given the chance to talk to a recruit for fifteen minutes, the lieutenant would be able to tell if he/she would be a quality officer.

## References

- Bartol, C.R. (1991). Predictive validation of the MMPI for officers who fail. *Psychology, Research and Practice*, 22: 127-132.
- Bayley, D., & Bittner, E. (1984). Learning the skills of policing. *Law and Contemporary Problems*, 47: 35-59.
- Buerger, M. (1994). A tale of two targets: Limitations of community anticrime actions. *Crime and Delinquency*, 40: 411-436.
- Burkhart, B. (1980). Conceptual issues in the development of police selection procedures. *Professional Psychology*, 121-129.
- Carter, D., & Radelet, L. (1999). *The Police and the Community*. New York: McMillan Press.
- Cordner, G. (1995). Community policing: Elements and effects. *Police Forum*, 5: 1-8.
- Doerner, W., & Hunter, R. (2006). Post-FTO performance evaluations of rookie officers. *Journal of Ethnicity in Criminal Justice*, 4: 113-128.
- Fyfe, J. (1999). Good policing. In Stojkovic, S, Klofas, J, & Klainich, D. (Eds.), *The Administration and Management of Criminal Justice Organizations*, 3<sup>rd</sup> Edition. Prospect Heights, IL: Waveland Press.
- Grant, J.D., & Grant, J. (1995). Officer selection and the prevention of abuse of force. In Geller, W., & Toch, H. (Eds.), *And Justice For All: Understanding and Controlling Police Abuse of Force*, Washington, D.C.: Police Executive Research Forum.
- Mastrofski, S. (1992). What does community policing mean for daily police work? *NIJ Journal* August:23-27.
- Metchik, E. (1999). An analysis of the “screening out” model of police officer selection. *Police Quarterly*, 2: 79-95.
- Muir, W. K. (1977). *Streetcorner Politicians*. Chicago, IL: University of Chicago Press.
- Parks, R., Mastrofski, S., DeJong, C., & Gray, M.K. (1999). How officers spend their time with the community. *Justice Quarterly*, 16: 483-518.
- Roberg, R., Kuykendall, J. C., & Novak, K. (2002). *Police Management*, 3<sup>rd</sup> Edition. Los Angeles, CA: Roxbury Publishing.
- Sanders, B. A. (2003). Maybe there’s no such thing as a “good cop”: Organizational challenges in selecting quality officers. *Policing: An International Journal of Police Strategies and Management*, 26: 313-328.

Smith, B., Novak, K., & Frank, J. (2001). Community policing and the work routines of street-level officers. *Criminal Justice Review*, 26: 17-37.

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

Travis, L. and Sanders, B. (1997) Community Policing Activities: The Ohio Task Analysis Project, Summary Report for the Ohio Office of Criminal Justice Services. Cincinnati: Division of Criminal Justice, University of Cincinnati.

Trojanowicz, R., & Bucquerox, B. (1990). *Community Policing: A Contemporary Perspective*. Cincinnati, OH: Anderson Publishing.

Walker, S., Alpert, G.P., & Kenney, D. (2000). Early warning systems for police: Concept, history and issues. *Police Quarterly*, 3: 132-152.

White, M. D. (2008). Identifying good cops early: Predicting recruit performance in the academy. *Police Quarterly*, 11: 27-49.