

University of Missouri, St. Louis

IRL @ UMSL

---

Dissertations

UMSL Graduate Works

---

2004

## Personality measurement in the prediction of positive and negative police officer performance

Brian P. Enright

Follow this and additional works at: <https://irl.umsl.edu/dissertation>



Part of the [Psychology Commons](#)

---

### Recommended Citation

Enright, Brian P., "Personality measurement in the prediction of positive and negative police officer performance" (2004). *Dissertations*. 1254.

<https://irl.umsl.edu/dissertation/1254>

This Dissertation is brought to you for free and open access by the UMSL Graduate Works at IRL @ UMSL. It has been accepted for inclusion in Dissertations by an authorized administrator of IRL @ UMSL. For more information, please contact [marvinh@umsl.edu](mailto:marvinh@umsl.edu).

Personality Measurement in the Prediction of Positive and Negative  
Police Officer Performance

Brian P. Enright

University of Missouri - St. Louis

A Dissertation in partial fulfillment of the requirements for the doctorate in  
Psychology.

## CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	vi
LIST OF FIGURES	vii
DEDICATION	viii
ABSTRACT	ix
<b>CHAPTER 1</b>	
INTRODUCTION	2
Psychological Assessment in Police Selection	2
Prosocial versus Antisocial Characteristics	7
Psychological Distress	16
METHODOLOGICAL ISSUES IN THE LITERATURE	19
Representative and Generalizable Samples	20
Gender and Ethnicity Factors	21
Theoretical Considerations in Police Selection Research	22
Objective versus Subjective Measures of Performance	25
Concurrent Validity versus Predictive Validity	29
MMPI and CPI in Police Selection	30

SUMMARY AND HYPOTHESES	33
Hypothesis One	35
Hypothesis Two	36
Alternate Hypotheses	36

## **CHAPTER 2**

METHOD	39
Participants	39
Measures	40
MMPI-2	40
CPI	43
Outcome Measures	45
Procedure	46

## **CHAPTER 3**

RESULTS	47
Preliminary Analyses	47
Data Transformations for Personality Measures	47
Data Transformations for Outcome Measures	49
Missing Data	49

Outliers	49
Homogeneity of Sample	50
Assumptions of Normality	52
Analyses and Modeling Strategy	53
Measurement Model	53
Structural Model	58
Follow-up Analyses	62
Additional Results	68

## CHAPTER 4

DISCUSSION	72
General Summary	72
Latent Constructs for Outcome	72
Latent Constructs for Personality Scales	74
Confound of Years of Experience	79
Objective versus Subjective Measures of Performance	83
Homogeneity of Sample	86
Conclusion	88

REFERENCES	91
------------	----

## Acknowledgments

I would like to again thank my committee members for their guidance. I would also like to thank all of my friends and family for their support and encouragement for the long period of time while this dissertation was in progress. Finally, this research could not have been possible without the generous cooperation and assistance of the Chiefs of Police from the two represented departments and their staff.

## List of Tables

Table 1. Mean, standard deviation, and range for all MMPI T scores.	43
Table 2. Mean, standard deviation, and range for all CPI scores.	44
Table 3. Mean, Standard Deviation, and Range for all Continuous Outcome Variables.	53
Table 4. K-corrected MMPI T-scores correlated with individual and composite outcome variables.	68
Table 5. CPI T-scores correlated with individual and composite outcome variables.	69
Table 6. Correlations between CPI scores and MMPI-2 T scores with and without the K-correction.	70
Table 7. Correlations between outcome variables and MMPI-2 T scores with and without the K-correction.	71

## List of Figures

Figure 1. Initial Measurement Model.	56
Figure 2. Modified Measurement Model.	60
Figure 3. Structural Model Testing the effect of psychological distress and prosocial characteristics on positive and negative outcome.	61
Figure 4. Structural Model Using Years of Experience as a Mediator Variable.	66
Figure 5. Simplified structural model with years-of-experience as a mediator variable using K-corrected T scores.	67



## Dedication

This dissertation is dedicated to all who work for peace.

## Abstract

Police officer selection research has revealed a number of relationships between scores on psychological tests and performance on the job. Although a few of these predictor scales have been consistently linked with job performance, many still require further investigation. It was proposed that the personality scales most consistently linked to job performance concern two broad constructs including: (1) *prosocial personality characteristics*, and (2) level of *psychological distress*. Guided by findings from the police outcomes literature and research regarding the factor structure and intercorrelation of MMPI and CPI scales, it was proposed that Scales 4 and 9 from the MMPI and Ac, Sc, Gi, Wb, and Re from the CPI represent a prosocial construct, while Scales F, K, 6, 7, and 8 from the MMPI represent a psychological distress construct.

For the present study, data from 20 female and 198 male police officers from two police departments were used to show the connection between pre-hire personality test scores and subsequent performance on the job. Structural equation modeling (SEM) was first used to evaluate for model fit and strength of relationship between the proposed latent variables and observed variables. Next, the structural model was tested to evaluate the hypothesized relationships between latent constructs. With some alteration to the proposed models, MMPI scales and CPI scales reliably loaded on latent constructs, but those constructs did not have

strong relationships with outcome variables in the predicted directions. Observed performance indicators reliably loaded on either a positive outcome latent variable or a negative outcome latent variable. The proposed prosocial latent variable required redefinition and showed a strong negative relationship with the positive outcome variable. The psychological distress factor did not show a strong relationship with either the positive or negative outcome latent variables.

Predicted and unexpected results are discussed in terms of the strengths and limitations of the current study and the use of measures of personality in the selection of police officer candidates.

## Personality Measurement in the Prediction of Positive and Negative

### Police Officer Performance

In 1967 the Presidential Commission on Law Enforcement and the Administration of Justice suggested that all police departments use psychological assessments as part of the hiring process for new police officers (Beutler, Nussbaum, Meredith, 1988). In that same year the National Advisory Commission on Criminal Justice Standards and Goals recommended that every police agency hire a psychiatrist or psychologist to test police applicants in order to rule out candidates who are emotionally unfit for police work (Burkhart, 1980; Meier, Farmer, Maxwell, 1987; Scogin & Beutler, 1986). Law enforcement agencies abide by these recommendations because of their sense of responsibility to the people they serve. Beyond a commanding officer's sense of duty to hire fit law enforcement personnel, there are also liability issues that provide incentive to utilize a rigorous selection procedure.

The doctrine of *respondeat superior* allows employers to be held legally liable for actions of employees (Costello, Schneider, Schoenfeld, 1996). Issues of negligent hiring or negligent retention extend beyond the behaviors of employees acting under official regulations; they also include reckless actions and other behaviors not consistent with authorized duties. For example, in a civil case (Hild

v. Bruner, 1980), it was concluded that a municipality's "...failure to conduct some kind of psychological tests of its police officers, at least after 1975 (when, according to expert testimony, such testing became widely accepted), constitutes gross negligence." (p. 99). The fact that an organization can be held liable for the actions of its employees even when those behaviors are in contradiction to organizational policy provides strong incentive to devote ample resources to the employee selection process.

Utilizing extensive background checks, physical abilities screening, and psychological evaluations in the process of hiring police officers invokes the sometimes conflicting responsibilities of government agencies to protect the rights of the individual and to ensure the safety of society as a whole. Although legal actions have challenged the constitutionality of conducting personality, cognitive, and even job-related evaluations (Pallone, 1992), in the case of safety-sensitive positions, the courts have consistently ruled in favor of protecting society as a whole and continuing to support the use of rigorous screening procedures. In a lengthy and pivotal court case (*McKenna v. Fargo*, 1978), the United States District Court in New Jersey established the permissibility to require job applicants to sensitive positions such as firefighting or policing to undergo psychological test-based evaluations to determine their fitness to serve. The court concluded, in part,

that "... (2) the applicant's right of privacy was burdened by the requirement, but (3) the interest of the City in screening out applicants who would not be able to handle the psychological pressures of the job was sufficient to justify the intrusion into the privacy on the applicant, ..." (p. 1355).

Ideally, in addition to being rigorous and thorough, these evaluations to determine fitness to serve should also be fair and consistent. *Griggs v. Duke Power Co.* (1974) was an influential case, placing an emphasis on evaluation instruments with proven predictive validity over more subjective methods of selection such as interviews alone (Kornfeld, 1995; Pallone, 1992). In 1986 the President of the International Association of Chiefs of Police (IACP) heralded psychological services as a vital component of a responsible department (Meier, Farmer & Maxwell, 1987). The IACP guidelines were recently revised to provide a comprehensive framework for psychological assessment of police candidates (Detrick, Chibnall, & Rosso, 2001). It is now standard practice for police departments and other agencies hiring for safety sensitive positions to utilize psychological assessment procedures in their personnel selection process.

Such measures have also been implemented as screening tools for other safety sensitive positions in the nuclear power and airline industries. Selection evaluations are also utilized for positions that require clearance to classified

material. Nuclear power employees are required to undergo regular evaluations to continually demonstrate emotional and psychological fitness for their position.

Although the intent to protect the community is sound, there are problems associated with repeated administrations of the same personality measure. Kelley, Jacobs, and Farr (1994) for example showed that with repeated administrations the MMPI, examinees scores became more and more normalized suggesting that practice effects provided the examinees opportunity to learn how to better answer the questions to give the impression of minimal distress. The authors suggest the use of the MMPI-2 for initial screening for employment, but more specific symptom focused measures for follow up screenings. The current study is focused exclusively on initial assessments administered to police officer applicants.

Other studies investigating the hiring of nuclear power employees and employees requiring high-level security clearance have been unsuccessful in using the MMPI in the prediction of outcome on various criterion measures (Muha & May, 1975; Wilikofsky, 1986). Butcher (1994) investigated the use of the MMPI-2 with airline pilot applicants and described the difficulties associated with interpretation of MMPI-2 profiles that fall largely within the normal range. He proposed the possibility of using specialized norm groups when using the MMPI-2 for selection purposes and tentatively concluded that it is reasonable to interpret

scale score elevations in the moderate range ( $T > 60$ ) when used for selection purposes. Overall, despite their use in a wide variety of safety sensitive industries, research regarding the use of personality measures in fields other than police officer selection is limited and rarely assesses future performance on the job after the hiring process.

Given their importance and pervasive use, the predictive utility of psychological assessments used in the selection of police applicants must be convincingly defensible. Although the research on most psychological assessment tools in general is extensive, research on the use of these tools in the selection of police candidates and the prediction of their future performance is comparatively limited and requires continual evolution. It is important that the research in this area continue to improve so that psychologists may provide this evaluation service effectively, both for the protection of the citizens in our communities and to be fair to the men and women who protect our safety.

Research of police performance has revealed a number of connections between scores on psychological tests and performance on the job. A few of these predictor scales have been consistently linked with job performance, while many still require further investigation. The personality scales most consistently linked to job performance appear to relate to one of two broad constructs including: (1)



prosocial personality characteristics, and (2) a candidate's level of psychological distress.

### *Prosocial versus Antisocial Characteristics*

Police work is more than a job from which to draw a salary, and it is not a means by which to seek one's fortune. Unlike self-employment, sales positions, careers in business, and many other jobs in the private sector, there is no promise or dream that with hard work and ingenuity officers can become rich in law enforcement. Giving up dreams they may have had of vast wealth when they don the badge, good police officers clearly have some other motivation than just material reward for pursuing their chosen career. Most police officers likely have prosocial characteristics such as a strong sense of duty and service to the community as motivating factors for their career choice.

As much as police work involves overt law enforcement activities, good police work also involves service to the public in more subtle ways. The duties of a police officer are overtly prosocial in that they include the capture of criminals, maintenance of the peace, and the assistance of persons in distress. However, more discrete aspects of good day-to-day police work are also related to a prosocial focus.

Most of the early research investigating prosocial behavior or an altruistic personality resulted in the conclusion that prosocial or helpful behaviors were largely situationally determined (Gergen, Gergen, & Meter, 1972). Some researchers asserted that efforts to investigate the altruistic personality leading up to the late 1970's had been fruitless, but research conducted in the 1980's showed more significant and replicable findings indicating that there are individual differences in willingness to help a person in distress (Davis, 1980; Rushton, Chrisjon, & Fekken, 1982). Although the initial findings were focused specifically on helping those in distress, the current state of research regarding prosocial personality traits seems to have a broader focus. Penner and colleagues (1995) indicated that they believe "prosocial behavior was too complex to be adequately predicted by a single personality characteristic (p. 148)." In their chapter on Measuring the Prosocial Personality, they investigated a group of existing personality scales previously found to be correlated with prosocial affects, cognitions, and actions. The actions frequently associated with prosocial personality include a wide variety of helping behaviors, taking responsibility, and honesty, with the most consistent emphasis placed on helping behaviors (Penner et al., 1995). For the purposes of the present study, the construct of prosocial traits is

used in an attempt to describe a broad range of characteristics and behaviors conventionally thought of as benefiting others.

As a personality characteristic, a community-oriented or prosocial focus should be predictive of good overall police officer performance. Although further clarification is needed, the research on police performance appears to support this assertion. Scales on psychological instruments that show a value for teaming, social interest, conformity, and rule-governed behavior predict a higher quality of job performance, whereas elevations on scales related to antisocial behavior, egocentrism, impulsiveness, and lack of consideration for others, seem to consistently predict problems in job performance (e.g., Bartol, 1991; Costello, et al., 1996; Hiatt & Hargrave, 1988; Hogan & Kurtines, 1975; Inwald & Susman, 1984).

Research using the Inwald Personality Inventory (IPI) indicates that prosocial characteristics lead to better performance in academy training. Police cadets who earned high scores on IPI scales linked to extroversion and agreeableness were significantly less likely to drop out of the academy, and they received higher overall training ratings compared to their counterparts who earned low scores on scales linked to extroversion and agreeableness (Cortina et al., 1992). Given the frequent need for collaborative interaction with fellow officers

and members of the public, it is not surprising that individuals who are low in agreeableness and extroversion receive lower ratings and are more likely to move on to other careers. The study by Cortina and colleagues (1992) is important in that it taps the positive end of the antisocial/prosocial spectrum; it is, however, a rare exception in this respect, particularly when the MMPI or IPI are used to measure personality.

The California Psychological Inventory (CPI) is a widely used measure of personality that adds important insights to this issue. The CPI contains a number of scales that tap favorable aspects of personality, including prosocial characteristics. CPI scales shown to have significant positive correlations with supervisory ratings include: Achievement via Conformance (Ac), Tolerance (To), Responsibility (Re), Self Control (Sc), Good Impression (Gi), Well-being (Wb), Achievement via Independence (Ai), and Independence (In) (Borstman, 1977; Hogan, 1971; Mills & Bohannon, 1980). The most frequently replicated finding is that for Achievement via Conformance, while Self-Control, Tolerance, and Responsibility each have been cited in more than one independent study as positively correlated with supervisory ratings.

The studies yielding these results in support of prosocial personality traits, however, are somewhat limited in that most utilize only supervisor ratings as

criterion measures, rather than including less subjective indicators of performance (e.g. suspensions from duty, written reprimands, or written commendations) or objective measures of outcome (e.g. motor vehicle accidents, injuries, or sick days). This and other methodological issues will be further discussed in subsequent sections of this paper. When methodological issues are considered, the Achievement via Conformance Scale is shown to be the most robust CPI scale for predicting police performance. Not only has Achievement via Conformance been frequently linked to positive supervisory ratings, but officers with high Ac scores were also demonstrated to have fewer disciplinary actions taken against them (Hogan & Kurtines, 1975).

Findings from the CPI clearly add to the IPI data supporting that prosocial characteristics lead to positive outcome. Respondents who earn high scores on Self-control and Responsibility generally are described as disciplined, diplomatic, nonjudgmental, and dependable (Gough, 1996; McAllister, 1988). Moreover, Achievement via Conformance is the scale most strongly and frequently associated with positive outcome, and those who earn high scores on this scale are rule-favoring, team-oriented, and generally work well with others (Gough, 1996; McAllister, 1988).

While CPI scales may be good tools for detecting prosocial characteristics, the existing police selection research using this measure is sparse and has limitations. The Minnesota Multiphasic Personality Inventory (MMPI), on the other hand, is represented in several methodologically sound studies and more directly taps the antisocial end of the spectrum.

The MMPI scales most strongly associated with antisocial personality characteristics are the Psychopathic Deviate (4) and Hypomania (9) Scales. People who earn high scores on these scales are often described as overactive, irresponsible, and untrustworthy. Their interpersonal relationships are shallow and superficial. They can create a strong initial impression because they are socially skilled and free from the anxiety, worry, and guilt that can lead to noticeable self-consciousness (Greene, 1991).

Two notable studies with sound methodology and sizable samples provide evidence that elevation on Scales 4 (Psychopathic Deviate) and 9 (Hypomania) generally predict future negative police performance. First, Bartol (1991) compared a group of 471 retained officers with a group of 44 who were fired or forced to resign. Results indicated that officers who earned high scores on a combination of K-corrected raw scores on Scales L, 4, and 9 were more likely to be terminated and rated as “immature” and “inappropriate” by supervisors. In fact,

his “Immaturity Index” (L+4+9) was used to correctly predict problem officers 74% of the time. If the immaturity index was exclusively used to make hiring decisions at the time of initial screening, however, 23% of the non-terminated officers would have been rejected, and 30% of the eventually terminated officers would have been accepted (Bartol, 1991). Although some correlations between personality measures and the outcome variable in this study were statistically significant, they indicated a small total effect size given that all correlation coefficients were less than .20.

The second study of note involved 107 subjects and used a combination of Scales F, 4, and 9 to predict the number of disciplinary suspension days administered to officers (Costello, et al., 1996). The combination of Scales F+4+9, known as the Husemann “aggression index,” was used to predict officers who accrued excessive disciplinary suspension days with 81% accuracy. However, 38% of non-problem officers also scored higher than the cut score (Costello, et al., 1996). The correlation coefficient of .223 between the combined score of the “aggression index” and the criterion indicated a low overall effect size. Although both of the previously described studies fall short of providing an index that can stand alone in the selection of police candidates, they each contribute to the research and indicate that Scales 4 and 9 of the MMPI can be used to predict future

negative job performance (Bartol, 1991; Costello, et al., 1996; Hargrave, Hiatt, & Gaffney, 1988).

Other studies using the MMPI to predict future outcome support the assertion that antisocial characteristics lead to poor police performance. Research focusing on experienced police officers revealed that serious disciplinary actions were more likely to have been taken against officers who had relatively higher scores on Scale 9 (Hypomania) and two other MMPI scales, Scales F and 6 (Hiatt & Hargrave, 1988). The relatively higher scores on the two other MMPI scales indicate the endorsement of unusual experiences and undue suspiciousness, which will be addressed further in the next section. The most prominent difference between the groups of officers in the Hiatt and Hargrave (1988) study occurred on Scale 9, where three times as many problem officers had an elevation as compared to non-problem officers. High scores on Scale 9 generally describe people who have difficulties with agitation, dissatisfaction, impulsiveness, ego inflation, and denial of guilt (Greene, 1991).

Another study that also used less subjective and objective criterion measures showed that Scale 4 elevations alone predicted police cadets who showed problems with absenteeism and who received significant disciplinary actions during academy training (Inwald & Susman, 1984). People who earn high



scores on Scale 4 frequently have problems with social conflict, problems with authority figures, and are egocentric (Greene, 1991).

IPI scales have been shown to tap a wide range of “acting out” or antisocial characteristics related to negative performance in police academy training (Inwald & Susman, 1984). For example, IPI scales Absence Abuse (a scale measuring frequent absenteeism), Drugs, Trouble With the Law, and Hyperactivity were related to problems with lateness. Elevations on scales labeled Job Difficulties, Substance Abuse, Drugs, and Driving Violations predicted problem absenteeism, while scales labeled Job Difficulties and Driving Violations predicted cadets who received at least one formal disciplinary action during training (Inwald & Susman, 1984). High scores on the Driving Violations scale also were shown to predict poor ratings and terminations in more experienced rural police officers (Mufson & Mufson, 1998). IPI scales such as Trouble-With-the-Law, Drugs, Driving Violations, and Job Difficulties, indicate a personality style that is focused on egocentric interests over prosocial concerns. These findings, therefore, provide further evidence that antisocial characteristics lead to negative job performance.

Previous research from diverse sources has demonstrated that a number of scales related to antisocial traits predict subsequent problems on the job. Though evidence for this effect comes from a variety of sources, Scales 4 and 9 of the

MMPI have been consistently implicated in future job difficulties. The Psychopathic Deviate (4) and Hypomania (9) Scales of the MMPI together tap such characteristics as social conflict, egocentrism, denial of guilt, and impulsiveness, which clearly describe antisocial characteristics.

### *Psychological Distress*

In addition to the prosocial construct, another broad category of characteristics related to failure in police work is an officer's level of psychological distress. Since good police work involves a unique set of skills and is often undertaken in highly stressful situations, it is important to select police candidates who are well adjusted and have good coping skills. Through the ordinary course of their workday, police officers expose themselves to danger and have the potential to confront life-and-death circumstances. It is, therefore, not only important to screen out those applicants who have easily diagnosable psychopathologies, but to select officers who have strong coping skills and will not be sensitive to high levels of stress (Beutler, et al., 1988; Burkhart, 1980; Scogin & Beutler, 1986).

The potential to be placed in emergency situations in police work is well known, and people's self-awareness regarding their own susceptibility to stress will often direct them toward other career fields. However, research shows that

officers who are not highly resilient to stress can have difficulties on the job, which indicates that the self-selection process is not completely effective (Beutler, et al., 1988; Cimbura, 2000; Riddle, 2000; Thomas-Riddle, 2000). A small sample of university police officers tested at recruitment, at two years post hiring, and at four years post hiring showed evidence of increased somatic complaints, anxiety, and alcohol vulnerability with greater experience on the job (Beutler et al., 1988). Although the authors warn against making broad generalizations based on their limited data, they do indicate that police officers can be susceptible to job stress even early in their careers.

Scales 6 (Paranoia) and 8 (Schizophrenia) of the MMPI are associated with psychological distress and have been used to predict negative performance in police officers. High scores on scales 6 and 8 describe individuals who express significant personal distress through their complaints of tension, worries, abnormal fears, difficulty concentrating, self-criticism, social alienation, anxiety, and depression. They are also seen as withdrawn, suspicious, and socially inappropriate (Greene, 1991).

High scores on Scales L (Lie), 6 (Paranoia) and 8 (Schizophrenia) were used together to predict poor performance on the job (Costello, et al., 1982). In fact, the Goldberg index  $[L+6(Pa)+8(Sc)-3(Hy)-7(Pt)]$  was used to classify officers

as acceptable or unacceptable with 74% accuracy. The authors were encouraged by their false positive rate of only 2%, but the Goldberg index successfully detected only 26% of the unacceptable officers (Costello, et al., 1982). Though this study used a predictive design and provides good information, some problems remain when using the Goldberg Index. First, it was implemented with little consideration given to its theoretical underpinning, which will be discussed further in a subsequent section of this paper. Secondly, the index incorporates scales loading in opposite directions, which normally have high positive correlations (i.e. Scales 7 and 8) (Hathaway & McKinley, 1989).

Other independent research has provided more evidence that problems with psychological distress are related to poor police performance. Costello and colleagues (1996) showed that high scores on the F Scale of the MMPI predicted disciplinary actions taken against officers, and Hiatt and Hargrave (1988) showed a similar outcome for those officers who earned high scores on Scale 6. While Scale 6 specifically taps distrustful and suspicious characteristics, the F Scale indicates unusual experiences in general and a high level of psychological distress.

Using MMPI and IPI scales as indicators of neuroticism, officers with high scores received poor peer ratings, poor instructor ratings, and made frequent use of counseling services (Cortina et al., 1992). Specific findings showed that the

MMPI Scale, Mac-R (Alcohol Addiction Potential) and IPI Scale Undue

Suspiciousness predicted police cadets who had problems with absenteeism during academy training. The Undue Suspiciousness Scale also predicted cadets who received at least one formal disciplinary action during training, and Unusual Experiences was related to problems with lateness (Inwald & Susman, 1984).

High scores on these MMPI and IPI scales generally indicate the endorsement of unusual experiences and the presence of psychological distress. These unusual perceptions of the environment and psychological difficulties would likely impair an officer's ability to cope with the inevitable stressors associated with police work. Because it is likely that police work can exact a considerable toll even for officers who are emotionally stable and have minimal psychological distress, it is imperative that people with lower resources for dealing with stress not be burdened with such responsibility.

### Methodological Issues

Many practical dilemmas emerge when attempting to investigate how well psychological measures administered to police applicants predict future performance on the job. Although existing studies have provided insight and direction, there are numerous methodological challenges in this type of research. Researchers in this area of study are confronted with difficulty obtaining large and

meaningful police officer samples, generating theoretically sound predictions, obtaining objective and less subjective measures of officer performance, utilizing predictive research designs rather than evaluating concurrent data, and utilizing comprehensive pre-hire measures of personality. Most of the previous research suffers from at least one of these limitations, thus reducing their robustness or generalizability.

### *Representative and Generalizable Samples*

Some previous researchers have had difficulty obtaining large and meaningful samples of police officers. Unless associated with a major metropolitan police department, it is difficult to obtain a significantly large sample of participants. Even if one can obtain a large sample of officers from a major metropolitan area, this sample may not be directly generalizable for use by the majority of police agencies in United States. Although the majority of police officers work for larger departments, approximately 96% of agencies employing sworn police officers have fewer than 100 officers and serve more suburban or rural communities (Hickman & Reaves, 2001). Although the research shows that there is some level of homogeneity within police officers' personality scores in general, there also appear to be significant within-group differences associated with size of department, type of community served, and the type of law

enforcement activities performed (Bartol, 1982; Carpenter & Raza, 1987; Hargrave & Hiatt, 1987; Hargrave, Hiatt, & Gaffney, 1986; Saccuzzo, et al., 1974; Saxe & Reiser, 1976). It is, therefore, important to consider these factors when designing a study. Participants should ideally be taken from departments of similar size that serve communities with similar demographic properties, and the results would be most applicable to other departments with similar demographics.

### *Gender and Ethnicity Factors*

Women and members of other minority groups are generally underrepresented in police officer selection and normative research. Two recent studies have contributed to an attempt to correct this under-representation. Kornfeld (1995) used the MMPI-2 to make gender and ethnicity comparisons with 61 male and 12 female Caucasian applicants, and 11 ethnic minority male police applicants. Detrick and colleagues (2001) conducted similar analyses using 395 male and 34 female Caucasian applicants, and 35 male and 3 female ethnic minority police applicants. Results from both studies were similar, indicating that regardless of race or gender, police applicants tended to present moderately defensive MMPI profiles “characterized by self confidence, lack of depression, interpersonal comfort, stereotyped male roles (for males), and rejection of traditional female roles (for women) (Detrick, et al., 2001) (p. 487).” Detrick and

colleagues (2001) found few statistically significant differences for ethnicity and combined, they only accounted for a small proportion of the total variability (<3%). No significant differences for gender were observed other than the expected finding for Scale 5 (Masculinity/Femininity). These studies are both limited by a small percentage of female and ethnic minority subjects, but the available data suggest there are more similarities than differences among male and female police applicants and applicants of diverse ethnicity. Although more research using larger samples of female and minority applicants is needed, for the present study data will be combined across gender and ethnic minority status for analysis.

### *Theoretical Considerations in Police Selection Research*

Recent studies have utilized previous MMPI findings to develop specialized indices for use in police selection, but not all of these indices were derived using sound theory regarding police officer's personalities. As described earlier, the Goldberg index  $[L+6(Pa)+8(Sc)-3(Hy)-7(Pt)]$  was used in such a study (Costello, et al., 1982). Despite its successes, some problems remain because it is difficult to determine what core personality features are being tapped by the Goldberg Index. It was originally empirically derived in the late 1960's to improve diagnostic classification of patients with psychotic disorders. The formula, which includes



high scores on the Lie (L), Paranoia (6), and Schizophrenia (8) scales and low scores on the Hysteria (3) and Psychasthenia (7), does not seem to pull together in a way that can helpfully describe aspects of non-psychotic personality. Therefore, it is difficult to use such an index to make inferences regarding specific personality characteristics of successful and unsuccessful police officers.

The concern is that without sound theoretical underpinning, we may find ourselves accepting or rejecting prospective officers for the wrong reasons. For example, Merian, Stefan, Schoenfeld and Kobos (1980) developed a five-item MMPI index empirically and without theoretical consideration. Although they reported some success with their original sample, the index was later found not to be generalizable to another police sample and the index was confounded with race (Wasson Dralle & Baybrook, 1985).

The “Immaturity Index” ( $L+4(Pd)+9(Ma)$ ), described earlier, is a multi-scale index from which theoretical implications were drawn based on interviews with supervisors involved in the study (Bartol, 1991). Bartol (1991) differentiated between acceptable and unacceptable police candidates by using his index to predict officers who were terminated or forced to resign. Although not comprehensive, the “Immaturity Index” is a somewhat successful predictor of negative police officer performance as indicated by the 74% rate of correct

detection though the correlations between predictor and outcome were modest. Moreover it was found to have theoretical underpinnings based on “immature” character traits. The scales that make up this index also seem to tap antisocial personality characteristics. As described earlier, the Psychopathic Deviate and Hypomania scales together tap such characteristics as social conflict, egocentrism, denial of guilt, and impulsiveness. The findings associated with high scores on the L (Lie) Scale, however, are less straightforward. Lie Scale elevations are typically associated with attempts to present oneself in an unrealistically favorable light, which is not necessarily consistent with either the construct of immaturity or antisocial tendencies. Moreover, other research has reported contradictory findings, indicating that low scores on the L Scale of the MMPI were linked to negative police performance (Hatt & Hargrave, 1988).

A more recent study, which made theory-based assertions regarding its multi-scale index, provides further support for the connection between Scales 4 and 9 and problems in job performance. This study, however, investigated the F (Infrequency) Scale rather than the L (Lie) Scale in conjunction with Scales 4 and 9 (Costello, et al., 1996). The combination of Scales F+4+9, known as the Husemann “aggression index,” was somewhat successfully used to predict disciplinary suspension days (Costello, et al., 1996). F Scale elevations, however,

seem more related to general psychological distress than aggressive behaviors, and may tap a separate factor than Scales 4 and 9.

Although this type of research may still be too young to determine a definitive raw score threshold that can be used in the absence of other information, the previously described studies have provided helpful information. This type of research has served to validate the use of the MMPI in police officer selection and has informed theoretical considerations regarding the personality characteristics of officers who fail. Scales 4, 9, and F are relatively reliable measures in the prediction of negative police outcome. One interpretation for elevations on Scales 4 and 9 is that an antisocial focus leads to poor police work, while F Scale elevations seem more related to a general psychological distress factor.

#### *Objective versus Subjective Measures of Performance*

As with measures of performance in other areas of psychological research, it is more valid, generalizable, and reliable to utilize objective measures of police performance over subjective measures. Previous research in this area has relied on a variety of outcome indicators including some objective indicators and various indicators that range in degree of subjectivity. Objective indicators include aspects of job performance that are easily operationalized and are unaffected by a supervisor's or peer's potentially biased opinion of an officer. Subjective

measures frequently used in this type of research include rating scales completed by an officer's supervisor or peers. Less subjective and more face valid measures that are used somewhat less frequently include significant recorded instances of disciplinary action or commendation.

Objective and less subjective measures provide important information regarding performance factors, yet they are somewhat limited in their scope. For example, these types of measures are much more easily and frequently obtained for negative indicators of performance (e.g. documented instances of disciplinary action, motor vehicle accidents, absenteeism, injuries, and terminations) than for positive performance. It is therefore a simpler task to identify the personality factors that predict poor or deficient police work than it is to predict the antecedents to outstanding future police work. Of the few objective and less subjective measures of positive performance available, most are inexorably confounded with issues such as opportunity bias. Opportunities to make arrests or issue traffic citations, for example, are at least in part dictated by an officer's shift, community, and duty assignment. Although some significant positive measures of performance are worth investigating (e.g. promotions and letters of commendation), they are often difficult to obtain and have a relatively low rate of occurrence. In past research, therefore, positive police performance has been

mostly the domain of highly subjective evaluations made by supervisors and, to a less frequent extent, peers.

Supervisor and peer ratings of performance suffer from problems with reliability and halo effects (Beutler, et al., 1985). Many police supervisors give overly positive evaluations to avoid having a negative impact on pay raises or promotions of the officers under their responsibility. Problems remain even if a researcher creates a behaviorally anchored rating scale and instructs the rating supervisor that scores will not be tied to officers' careers. There is some evidence that when presented with detailed specific descriptors of various performance domains, police supervisors often ignore dimensions of the scale and, instead, utilize their overall impression of the officer being evaluated (Fitzgerald , 1987).

In addition to the obvious drawbacks of supervisor ratings, including ceiling effects and limited variability, the research investigating these subjective measures shows further limitations. Supervisor ratings of police performance have not been consistently linked to the more face valid and objective outcome indicators of performance. For example, Cascio and Valenzi (1978) compared several objective or performance based indices to supervisor ratings obtained using Behaviorally Anchored Rating Scales (BARS). They found that less than 25% of the variance in the supervisory ratings was accounted for by the performance based and

objective indices. Moreover, Beutler and colleagues (1985) found some scales were associated with opposing results on rating scale based measures and less subjective indicators of performance. Specifically, they observed a positive correlation between high supervisor ratings and elevated scores on Scale 2 (Depression) of the MMPI, and a significant negative correlation between the same MMPI scale and less subjective measures of good performance (i.e. low number of citizen grievances and absence of suspensions). The cumulative effect of these findings raises “doubt about the value of supervisor ratings as criteria for validating selection procedures” (Beutler, et al., 1988, p. 504).

Although the reliability of supervisory ratings can be enhanced with training, they are still plagued by a number of methodological problems (Bartol, 1991). There are clear drawbacks to using subjective ratings, but they have been frequently reported as the primary or sole criterion measure. Adding to the unavoidable drawbacks faced by researchers using supervisor ratings, some studies that have relied upon these ratings have used only a single rater or failed to conduct tests of inter-rater reliability (e.g. Cortina, et al., 1992), thereby having no means by which to evaluate the consistency of their criterion measure, much less to make inferences regarding its validity. There are, however, some notable exceptions to this problem in the research. Noting the limitations of using

supervisor ratings, and despite his many years of longitudinal subjective data, Bartol (1991) elected to report the less subjective and more face valid measure of terminations rather than supervisory ratings as his primary finding.

The type of criterion measure utilized is a crucial issue when evaluating research in the police selection literature. Therefore, the type of performance criteria reported was closely scrutinized for the current review, and the present study exclusively relied upon objective or less subjective outcome measures rather than utilizing supervisor ratings.

#### *Concurrent Validity versus Predictive Validity*

Concurrent validity research is characterized by analyzing data where test scores are obtained at approximately the same point in time as the criterion measures (Cohen, Swerdlik & Smith, 1992). An example of a concurrent design would be if researchers administered personality measures to several active duty police officers, and around the same time asked for supervisor ratings of all of the same officers' performance. Concurrent validity, therefore, can be used to estimate the relationship between an individual's test score and the outcome on a particular criterion measure. Predictive validity measures the relationship between test scores and a criterion measure obtained at a subsequent point in time (Cohen, et al., 1992). An example of a predictive design would be if several police officers

were administered personality measures prior to joining the force, then supervisors were asked to rate each of the officers after three years of experience. Predictive validity research designs can show how accurately scores on a pre-hire battery of tests (e.g., MMPI and CPI) predict future performance on criterion measures (e.g., supervisor rating, number of suspension days, or number of commendations). A predictive validity design, therefore, is the preferred method for validating a police officer screening procedure. The time and expense involved in utilizing a predictive design, however, limits some researchers to conducting concurrent validity studies.

#### *MMPI and CPI in Police Selection*

One of the limitations preventing more reliable indices for police selection may be the almost exclusive reliance on MMPI scales in this effort. The primary purpose of the MMPI is to evaluate level of functioning in potential psychiatric patients. It, therefore, has been successfully used in the practice of police officer selection by screening out applicants with significant psychopathology. More research, however, is still needed regarding the utility of the MMPI in personnel selection or for the prediction of specific outcome measures. Although there is some promising research using the MMPI in personnel selection, since the test's primary purpose is to detect significant psychopathology it should be augmented



with other measures in the selection of police officer candidates. A more comprehensive test battery may provide greater insight into candidate personality traits, thereby allowing for more measures from which to predict performance. In an early study, Saxe and Reiser (1976) illustrated an important issue regarding the use of MMPI basic scales in police officer selection that still has relevance.

Although significant differences were noted between the groups studied, these differences were “all within the ‘normal’ range and are too small in terms of traditional clinical standard scores to have meaningful utility in clinical differentiation of successful and unsuccessful police applicants” (Saxe & Reiser, 1978, p. 424).

The MMPI was not created to differentiate traits in normal or high functioning people. Rather, it was conceived from a pathology model and empirically derived using clinically psychopathological criterion groups. It should, therefore, not be surprising that a problem of range restriction exists when studying police officer candidates using this measure (Bernstein, Schoenfeld, & Costello, 1982). In fact, some researchers believe the MMPI should exclusively be used to measure maladjustment rather than personality in general. Butcher and Tellegen (1978), for example, argue that the MMPI is not sensitive to “normal

range” personality attributes and that alternative personality measures should be used in nonpsychiatric settings.

Despite arguments against the use of the MMPI in nonpsychiatric settings, the MMPI continues to be used widely in such situations and researchers have investigated possible interpretations of MMPI data in nonclinical populations (Kunze & Anderson, 1984). Further research in this area, therefore, may provide direction for clinicians who already utilize the MMPI to screen out applicants with significant psychopathology in terms of subtle and job-specific predictions of outcome. Some research investigating sub-clinical range MMPI scores has already provided support for this effort. Butcher (1994) studied the use of the MMPI in the selection of airline pilots and found results in support of using interpretations of moderate T score elevations ( $T > 60$ ) in cases of personnel selection. The MMPI’s limitations with normal populations nonetheless point to other complimentary psychological measures that may improve prediction of positive and negative officer performance. The California Psychological Inventory (CPI) is a well-validated personality measure ideally suited for coupling with the MMPI in prediction of police performance. This measure can be used to detect personality traits on the higher end of the continuum of functioning, including the presence of

positive and prosocial personality traits, while the MMPI is best suited for detecting the antisocial end of the spectrum and detecting psychological distress.

Compared to the MMPI, much less research has been conducted using the CPI in relationship to police officer performance despite its potential to improve the prediction of success and failure on the job. In general samples, the psychometric properties of the CPI hold up well in comparison to other widely used personality measures. For example, test-retest and internal consistency compare well with other measures of personality (Anastasi, 1988). Unfortunately, most of the previous research using the CPI to describe or predict police officer performance contains drawbacks that limit potential conclusions. Most studies suffer from one or more of the following limitations: use of a limited sample, primary reliance on subjective measures of performance (e.g., Borstman, 1977; Hogan, 1971; Mills & Bohannon, 1980; and Pugh, 1985), use of a concurrent design (e.g., Hogan, 1971; Hogan & Kurtines, 1975), and nearly all use CPI scales alone as predictors rather than incorporating scales from a comprehensive test battery.

### Summary and Hypotheses

Some promising trends and other troubling gaps exist in the literature regarding psychological assessment and the prediction of police officer outcome.

The MMPI has been extensively utilized in this type of research and some frequently replicated findings have emerged. Most notably, elevations on Scales 4 and 9, which are related to antisocial and egocentric characteristics, were found to predict a number of serious problems in performance. Scales 6 and 8 of the MMPI have, somewhat less reliably been linked to negative police outcome.

The CPI may be a useful complement to the MMPI with the population of interest due to the problem of range restriction when the MMPI is used in personnel selection, but the existing research using the CPI does not reveal many frequently replicated findings and lacks strong validation. One notable exception is the relationship between high scores on Achievement via Conformance and success on the job. Although less reliably, scores on Self-control and Responsibility have been linked to positive outcome indicators. Despite some replicated findings, much of the research in this area, whether using the CPI, the MMPI, or both, is plagued by one or more methodological problems.

Significant practical dilemmas emerge when investigating psychological measures and their predictive ability of future police officer performance: 1) The majority of police departments are relatively small, resulting in many studies with small sample sizes, 2) the difficulty involved in obtaining objective or face valid criterion measures of performance led many researchers to rely heavily on less

valid and reliable subjective ratings as measures of performance, 3) the time involved in employing longitudinal studies has led some researchers to employ concurrent rather than predictive designs, and 4) the ubiquity of the MMPI in police screening procedures has contributed to a paucity of research investigating other personality measures. The current study was developed to minimize or eliminate many of the failings of previous research by gathering data from a large sample of participants, utilizing a predictive validity design, employing objective or less subjective measures of officer performance, and using scales from both the MMPI and CPI as predictors.

For the current study, police officer outcome research, factor analytic research of the MMPI and CPI (Gough, 1996; Hathaway & McKinley, 1989; McAllister, 1996; Mitchell & Pierce-Jones, 1960; Nichols & Schnell, 1963) and theoretical considerations were used to derive the following hypotheses:

1. Scores on MMPI-2 and CPI scales hypothesized to represent prosocial characteristics will be associated with positive performance, while scores in the converse direction, which represent antisocial characteristics, will be associated with negative performance. MMPI-2 scales 4 (Psychopathic Deviate) and 9 (Hypomania), and CPI scales Ac (Achievement via Conformance), Sc (Self Control), Gi (Good Impression), Wb (Well-being), and Re (Responsibility) are

predicted to load on the same latent construct representing a continuum of antisocial to prosocial personality characteristics where the two MMPI-2 scales are scored in the opposite direction from the CPI scales. It is further predicted that this resulting latent construct will be positively correlated with good police officer performance and negatively correlated with poor police officer outcome.

2. High scores on personality scales associated with high psychological distress will show an association with poor performance while scores in the converse direction, which represent minimal psychological distress, will be associated with positive performance. MMPI-2 scales 6 (Paranoia), 7 (Psychesthenia), 8 (Schizophrenia), F (Infrequency), and K (Correction) from the MMPI-2, are predicted to load on a single latent construct representing a continuum of high to low psychological distress where the K scale is scored in the opposite direction from the other scores. It is further predicted that this latent construct will be negatively correlated with good police officer performance and positively correlated with poor police officer outcome.

### *Alternate Hypotheses*

The following alternatives to the previously proposed model are offered in the event that the previously proposed model does not provide a good fit of the data. Information relevant to the research question was gathered from diverse

sources and a number of possible predictions could have been made based on available data. Although there is reason to believe that the F and K Scales of the MMPI may be significantly associated with the psychological distress factor, previous correlational research has shown them both to be highly associated with scales that are included in the proposed prosocial factor. High correlations with several scales from both proposed latent variables would obscure the unique contribution of each latent variable. Therefore, one possible alteration would be to simply eliminate the F and K Scales from the model.

MMPI-2 scales Pd and Ma have been linked in previous research with the prediction of officer outcome, and elevations of these scores could be theoretically thought of as representing antisocial traits. However, Pd also has been shown to have high correlations with MMPI scales chosen to represent the psychological distress factor (i.e. Sc and Pt), and low scores Pd and Ma may simply be representative of the absence of antisocial personality traits and are not necessarily associated with prosocial personality traits. Therefore, other possible alterations to the model could include attaching the Pd scale in the psychological distress factor rather than the prosocial factor, remove Pd and Ma from the prosocial factor and set them to load on a separate antisocial factor, or simply eliminate Pd and Ma from the model.

Although it is hypothesized that all outcome measures may not load on a single latent construct, they are all set to load on one construct in the initial proposed measurement model. It is hoped that all outcome measures load on a single latent variable thereby reducing measurement error and increasing the reliability of the outcome latent. An alternate hypothesized result, however, is that the negative outcome measures (i.e. FIRE, MVA, REP, and SUS) will load on a single negative outcome latent while the positive outcome variables (i.e. CMD, CMM, and PRO) will load together on a separate latent.



## Method

### *Participants*

Analysis included data from 20 female officers and 198 male officers who were administered a standardized pre-employment evaluation as part of the hiring process, and who were subsequently accepted into their respective police departments. Participants included in the study underwent pre-employment screening during a fifteen-year window between 1985 and 2000. All female officers were Caucasian and only four of the male officers were people of color, while the remaining male participants were Caucasian. All participants were employed by one of two Midwestern police departments of similar size. Both departments employed between 70 and 90 sworn police officers for the inclusive years in which data were collected. Both departments serve a medium-sized suburban population consisting of middle to upper-middle class residents. Participants used in the current analysis included 104 cases from Department One, and 114 cases from Department Two. Information from both active duty and retired officers is included. Demographic information obtained for each subject included the following: sex, age, race, type and extent of previous experience, type and level of education, previous volunteer experience, and current rank.

*Measures*

Although the pre-employment battery includes other measures, the primary assessment tools used in this study are the MMPI and CPI.

1. The Minnesota Multiphasic Personality Inventory –2 (MMPI-2)

(Hathaway & McKinley, 1989) is a 567-item, true-false, objective, and empirically derived instrument designed to assess personality and emotional disorders. The normative sample of the revised version consists of 1,138 males and 1,462 females. The MMPI is currently the most widely used and researched objective personality inventory (Greene, 1991). The MMPI-2 includes the following four validity scales and 10 basic Scales: ? (Cannot say), L (Lie), F (Infrequency), K (Correction), 1 (Hypochondriasis), 2 (Depression), 3 (Hysteria), 4 (Psychopathic Deviate), 5 (Masculinity-Femininity), 6 (Paranoia), 7 (Psychasthenia), 8 (Schizophrenia), 9 (Hypomania), and 0 (Social Introversion). Scales are computed and profiled to yield standard T-scores. Retest reliability coefficients at a one-week interval, range from .67 for Pa to .92 for Si for adult males and from .58 for Pa to .91 for Si for adult females. Retest coefficients at a one week interval for the scales used in the current study range from .67 for Pa to .89 for Pt for males and from .58 for Pa to .88 for Pt for females (Hathaway & McKinley, 1989).

All of the MMPI-2 clinical scales used in the current study were empirically derived and validated comparing the item endorsements of patients from criterion groups to responses from subjects in normative samples. The Pd Scale was constructed and validated empirically using criterion groups of individuals diagnosed as psychopathic personality. The Pa scale was empirically derived by contrasting the item responses of a relatively heterogeneous group of paranoid patients with the original normative sample. The Pt scale was developed and validated comparing criterion groups of patients with abnormal fears, compulsive rituals, obsessions, difficulty concentrating, and self-criticism with the normative sample. The Sc scale was empirically derived and validated by comparing item responses from samples of patients diagnosed with schizophrenia to the normative sample. The Ma scale was empirically derived by comparing the item responses of patients suffering from mild to moderate symptoms of mania to the original normative sample (Greene, 1991).

The validity scales included in the current study were derived to assess for various issues that may affect the validity of individual test administrations. The F scale was originally derived to detect unusual or atypical ways of answering the test items and is comprised of items that no more than 10 percent of a normative sub-sample answered in the deviant direction. The K scale consists of items that

were empirically selected to identify patients with significant psychopathology yet had clinical scale profiles in the normal range (Greene, 1991). Primary analyses in the current study utilized the K-corrected T scores for MMPI-2 scales. This decision was made based on multiple factors: 1) the hypotheses for the current study were generated based on previous research utilizing K-corrected T scores, 2) since K-corrected T scores have typically been used in this type of research in the past, the current findings were thought to be more comparable to the existing literature if K-corrected T scores were examined, and 3) since the majority of clinicians utilizing the MMPI-2 to evaluate police officer candidates typically evaluate K-corrected T scores for basic scales, the current results were thought to be more directly clinically applicable if K-corrected T scores were evaluated. It should be noted, however, that there is an evolving literature indicating that using non-K-corrected T scores rather than K-corrected T scores may be warranted in the case of personnel selection or any other circumstance when the MMPI-2 is used to evaluate a largely psychologically healthy population (Barthlow, et al., 2002; Detrick, et al., 2001).

Variable	Mean	Standard Deviation	Range
F	43.61	5.04	33 – 57
K	61.60	8.10	37 – 77
Pd (4)*	49.87	6.75	33 – 69
Pa (6)	47.01	7.55	30 – 64
Pt (7)*	46.08	5.87	30 – 60
Sc (8)*	46.20	5.62	31 – 63
Ma (9)*	47.92	6.74	34 – 69

Table 1. Mean, standard deviation, and range for all MMPI T scores used in the current study (N = 218). \*K-correction applied for Pd, Pt, Sc, and Ma.

2. The California Psychological Inventory (CPI) is a 434-item objective personality measure constructed to assess individuals by means of “folk measures,” which are proposed to be universal and relevant to everyday life. The folk concepts are easily understood by ordinary people and are designed to capture the uniqueness of individuals in their social interactions. The CPI scales are meant to assist in the understanding, classifying, and predicting of individual behavior.

Scores are computed and profiled to yield standard T-scores for 20 folk scales. Of the 20 folk scales, the following five will be included in this study: Responsibility (Re), Self-control (Sc), Good Impression (Gi), Well-being (Wb), and Achievement via Conformance (Ac). The normative sample of Form 434 consists of 3000 males and 3000 females.

Regarding the reliability of Form 434, alpha coefficients for the scales used in the current study for males range from .77 for Re to .84 for Wb. Alpha coefficients for females range from .76 for Re to .85 for Wb. Test-retest correlations across a 25-year interval for adult males range from .49 for Gi to .82 for Do. For adult females, 5-year test-retest correlations range from .36 for Ac to .73 for Sc. One-year test-retest correlations for high school females range from .68 for Gi to .78 for Ac. For all 20 folk scales, the median one-year test-retest correlation for high school students was .68. For the scales used in this study, the median one-year test-retest correlation for high school students was also .68. The CPI has been extensively cross-validated with a variety of other measures including intellectual functioning, personality, academic performance, and personal observation. Data regarding the extensive cross validation information is contained in the 1996 CPI manual (Gough, 1996).

Variable	Mean	Standard Deviation	Range
Re	56.42	7.05	35 – 72
Sc	58.96	7.88	35 – 76
Gi	61.40	10.82	35 – 83
Wb	59.91	5.17	44 – 76
Ac	61.97	4.98	46 – 71

Table 2. Mean, standard deviation, and range for all CPI scores used in the current study (N = 218).

3. Outcome measures used in the analysis included negative and positive instances of officer performance. These indicators were obtained from officers' personnel files and include the following:

*Motor vehicle accidents (MVA).* Accidents were recorded if they occurred on the job while the patrol vehicle was in motion and where damage was caused. Only the accidents that were deemed avoidable or due to driver error were included in the analysis.

*Written reprimands (REP).* Any formal written reprimand that has become a permanent part of an officer's personnel file was included in the analysis.

*Suspensions from duty (SUS).* The total number of days an officer has been suspended from duty was included in the analysis.

*Terminations or firings (FIRE).* All officers unilaterally forced to leave their position were classified as a termination. The following are inclusive in this category: official terminations, forced resignations for disciplinary reasons, and failed probationary periods.

*Written commendations (CMD).* The total number of significant recognitions for outstanding performance from the department or from another policing agency was included in the analysis as a measure of positive officer performance.

*Major commendations (CMM).* The total number of “officer of the year” awards and national commendations for each officer was included in the analysis as a measure of positive officer performance.

*Promotions (PRO).* All promotions above the rank of Patrol Officer were included in the analysis as a measure of positive officer performance.

### *Procedure*

Participants were administered the MMPI and CPI as part of a battery of psychological assessment measures used to evaluate police officer candidates. The psychological evaluation represented one aspect of a comprehensive screening process used by the hiring managers of each department.

Two police departments participated in the study. Participation consisted of making officers’ personnel files available to the researcher so that performance data could be collected and recorded. All data collection using department personnel files was conducted on department premises. Participants used in the current analysis included 104 officers from Department One, and 114 officers from Department Two.

All gathered personnel information were used only for the purposes of these analyses and were kept strictly confidential. Only group data are reported and individual cases were assigned a number.



## Results

### *Preliminary Analyses*

*Data transformations for personality measures.* Due to the longitudinal nature of this research, data from both the current forms (i.e. MMPI-2 and CPI 434) and the earlier versions (i.e. MMPI and CPI 462) of personality measures were used in the current study. The MMPI-2 represents the re-standardization of the MMPI that served to provide current norms, a larger and nationally representative sample, include appropriate representation of minority groups, and update item content where needed. Despite the need for modification, “the items on the validity and clinical scales of the MMPI are essentially unchanged on the MMPI-2 except for the elimination of 13 items based on item content and the rewording of 68 items (Greene, 1991) (p. 20).” Research suggests that the MMPI-2 functions very similarly to the MMPI both at the item level and at the relative scale level. “In fact, with regard to relative scale standing and codetypes, the MMPI-2 matches closely enough to be considered an alternate form of the same test (Weed & Butcher, 1992) (p. 148).” Though Weed and Butcher (1992) reported that absolute differences between the MMPI and MMPI-2 clinical scale elevations can be “summarized by considering the MMPI-2 scale elevations to be approximately five T-score points lower than the corresponding MMPI scale” (p.

148), in the present study a more precise comparison of MMPI to MMPI-2 scores was used. The T-score elevations from the MMPI compared to the MMPI-2 are not consistently five points for all subscales. Since the item content and length remained identical or very similar from MMPI to MMPI-2 for most subscales, raw scores were used to more accurately transform MMPI scores to estimated MMPI-2 scores. All scale scores on the original MMPI subscales were converted to raw scores, which in turn, were converted into T-scores for the MMPI-2 using table K.1 in Appendix K of the Hathaway and McKinley MMPI-2 Manual for Administration and Scoring (1989). Due to a slightly shorter F scale in the MMPI-2, a prorating procedure was conducted to convert MMPI raw scores to MMPI-2 raw scores prior to converting back into T-scores for the MMPI-2.

Data from both the CPI 434 and the earlier version (CPI Form 462) are used in the current study. Form 434 of the CPI contains 28 fewer items than the previous version (Form 462). These 28 items were eliminated due to the 1991 Americans with Disabilities Act, fair employment legislation and privacy rights. With regard to comparison of Form 462 with Form 434, Gough (1996) reported that correlations between the 462-item and 434-item versions of the scales ranged from a low of .96 for Sa and Wb to a high of 1.0 on six scales that are scored the same way on both forms. Differences between scale means, for males, ranged

from 0 to 1.11 for Wb with an overall mean difference of 0.24. Differences between scale means, for females, ranged from 0 to 1.50 for Wb with an overall mean difference of 0.29. Gough (1996) cited the minimal alteration between forms, the high correlations between scales of the two forms, and the low magnitude of differences in score means to suggest that the comparability of the two forms has been maintained. For the present study, scale scores from both versions of the CPI were used without transformation.

*Data transformations for outcome measures.* Outcome data were recorded as accumulated incidences, which resulted in officers with more experience on the force at the time of data collection having more opportunity to accrue both positive and negative indicators of performance. Scores for negative and positive outcome were, therefore, divided by each individual participant's number of years of experience to adjust for years of service.

*Missing data.* Cases with missing MMPI or CPI data were eliminated from the analysis.

*Outliers.* Univariate outliers were changed to a value that was three standard deviations away from the mean for that variable (Kline, 1998). This procedure was conducted for a total of 17 observations across three outcome variables, which represents approximately 01.0% of all the outcome variables.

*Homogeneity of sample.* Significance tests were conducted to assess for differences between samples from the two police departments. Chi-Square tests revealed no significant differences between the two departments for the following demographic information: sex, pre-hire inner city policing experience, pre-hire rural policing experience, pre hire law enforcement volunteer experience, level of education, type of education (i.e. law enforcement related degree or other), and present rank. A significant difference was found for pre-hire suburban policing experience ( $\chi^2 = 5.51$ ,  $df = 1$ ,  $p < .01$ ). Regarding pre-hire suburban policing experience, 41.6% of participants from Department Two had previous experience serving in a suburban police department, whereas only 38.4% of participants from Department One had previous suburban police department experience. Although significance tests could not be conducted for race due the low frequency of group membership, it should be noted that the sample from Department Two included four minority participants while there were no minority participants in the sample from Department One.

An independent samples t-test revealed a statistically significant difference in participant age at  $p < .01$ . The mean age for participants from Department One was 36.01 and the mean age for participants from Department Two was 40.57. An independent samples t-test also showed a statistically significant difference in

years of experience in the department prior to the collection of follow up data at  $p < .01$ . The mean years of experience for participants from Department One was 5.72 and the mean years of experience for participants from Department Two was 8.02. The effect of the resulting mean difference of 2.29 years is likely minimized due to the previously described data transformation applied to reduce the impact of years of service. Although the majority of demographic factors measured were not significantly different between the two departments, the mean difference in age of 4.53 years between the two sub-samples may be evidence that the two samples are not representative of the same population. This issue was considered when interpreting the results of the analyses to follow.

Concerns regarding the possible differences in frequency of positive and negative outcome variables between the two departments prompted a comparison of these variables between Department One and Department Two. Independent samples t-tests comparing the continuous outcome variables between the two departments were significant for CMD at  $p < .05$ , CMM at  $p < .05$ , and REP at  $p < .05$ . Means for CMD were .486 for Department One and .635 for Department Two; for CMM were .047 for Department One and .019 Department Two; for REP were .068 for Department One and .164 for Department Two. Although these differences were statistically significant, the small actual mean difference in scores

indicates that there may be low practical significance. Nevertheless, these findings may have implications regarding generalizing findings across departments and they were considered when interpreting the analyses to follow.

*Assumptions of normality.* The distributions of all variables were evaluated for normality. As expected, scores on standardized personality measures were normally distributed. Several outcome variables, however, were positively skewed, including number of days suspended from duty (SUS), written reprimands (REP), motor vehicle accidents (MVA), written commendations (CMD), major awards (CMM), and promotions (PRO). Data transformations were completed by taking the square root of the skewed variable, which resulted in reducing skew and kurtosis. Although this transformation did not result in completely normal distributions for all variables, there is some evidence that the maximum likelihood estimation method used in the current SEM analyses is robust to the normality assumption (Hayduk, 1987; Joreskog & Sorbom, 1989). Descriptive statistics for the continuous outcome variables can be found in Table 1. Regarding frequencies for the one dichotomous outcome variable (FIRE), the sample included 200 retained officers and 18 officers who were fired or forced to resign.

Variable	Mean	Standard Deviation	Range
MVA	.1756	.2576	0 – 1.0
REP	.1907	.2878	0 – 1.73
SUS	.0895	.2599	0 – 1.66
CMD	.5864	.4568	0 – 1.73
CMM	.0254	.0911	0 - .50
PRO	.0385	.1125	0 - .58

Table 3. Mean, standard deviation, and range for all continuous outcome variables.

*Analysis and Modeling Strategy*

Structural equation modeling (SEM) was used to examine hypothesized interrelationships among latent constructs. Maximum likelihood (ML) methods were used to estimate model parameters and analyses were conducted using Amos Version 4.0 (Arbuckle & Wothke, 1999). The path model of hypothesized relationships was tested using a two-step modeling approach involving first estimating the measurement model followed by the structural model (Anderson & Gerbing, 1988, Schumacker & Lomax, 1996). In the first step, the measurement model was evaluated for model fit and strength of relationship between proposed latent and observed variables. The second step involved testing the structural model: that is, the hypothesized relationships between latent constructs.

*Measurement Model.* The initial measurement model is shown in Figure 1. In the initial model, the prosocial latent factor was represented by seven indicator

variables: Pd and Ma scales taken from the MMPI-2 and coded in the negative direction, and the WB, Gi, Ac, Re, and Sc scales from the CPI. The psychological distress latent variable was comprised of five indicator variables: Pa, Pt, Sc, F, and K were all taken from the MMPI-2, and K was coded in the negative direction. The latent factor, police officer performance, was represented by seven indicator variables: reprimands (REP), suspension days (SUS), at fault motor vehicle accidents (MVA), firings (FIRE), which were all coded in the negative direction, and commendations (CMD), major commendations and awards (CMM), and promotions (PRO). All of these variables were continuous with the exception of FIRE, which was entered as dichotomous with a 1.0 indicating a firing or a failed probationary period and a 0.0 indicating officer retention.

To assist in interpretation of the model, when possible, variables were transformed so that high scores could be interpreted in an intuitive direction based on the name of the latent variable. Each indicator variable had an error term with a standardized loading set to 1.0. In addition, each factor was scaled by setting the loading of one of its indicator variables to 1.0. This standardizes the latent variable by giving it the same measurement scale as the indicator variable. In the first step of the analyses, the initial measurement model was tested by allowing all latent variables to covary with each other.



A combination of traditional ML chi-square ( $\chi^2$ ) goodness-of-fit statistics was selected to evaluate the overall ability of the model to fit the data: relative chi-square ( $\chi^2/\text{df}$ ), root-mean-square error of approximation (RMSEA), Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI). Good model fit was considered achieved if the value of ( $\chi^2/\text{df}$ ) < 2.00, (RMSEA) < .05, (CFI) > .90, and (TLI) > .90 (Arbuckle & Wothke, 1999).

Fit indices of the initial measurement model represented an inadequate fit ( $\chi^2 = 611.37$ ,  $\text{df} = 149$ ;  $\chi^2/\text{df} = 4.10$ ; RMSEA = .120; CFI = .727; TLI = .687), and the standardized parameter estimates for many indicator-latent relationships were low (See Figure 1). The initial model was therefore modified until an acceptable fit was attained.

Consistent with pre-hoc hypotheses regarding how the model may need revision, standardized loadings for the MMPI-2 subscales Pd and Ma on the prosocial factor were quite low (absolute values of less than .21), and Pd had a negative parameter estimate indicating that it may be more appropriately attached to the psychological distress latent factor. Therefore, Ma was eliminated from the model and Pd was set to load on the psychological distress factor.

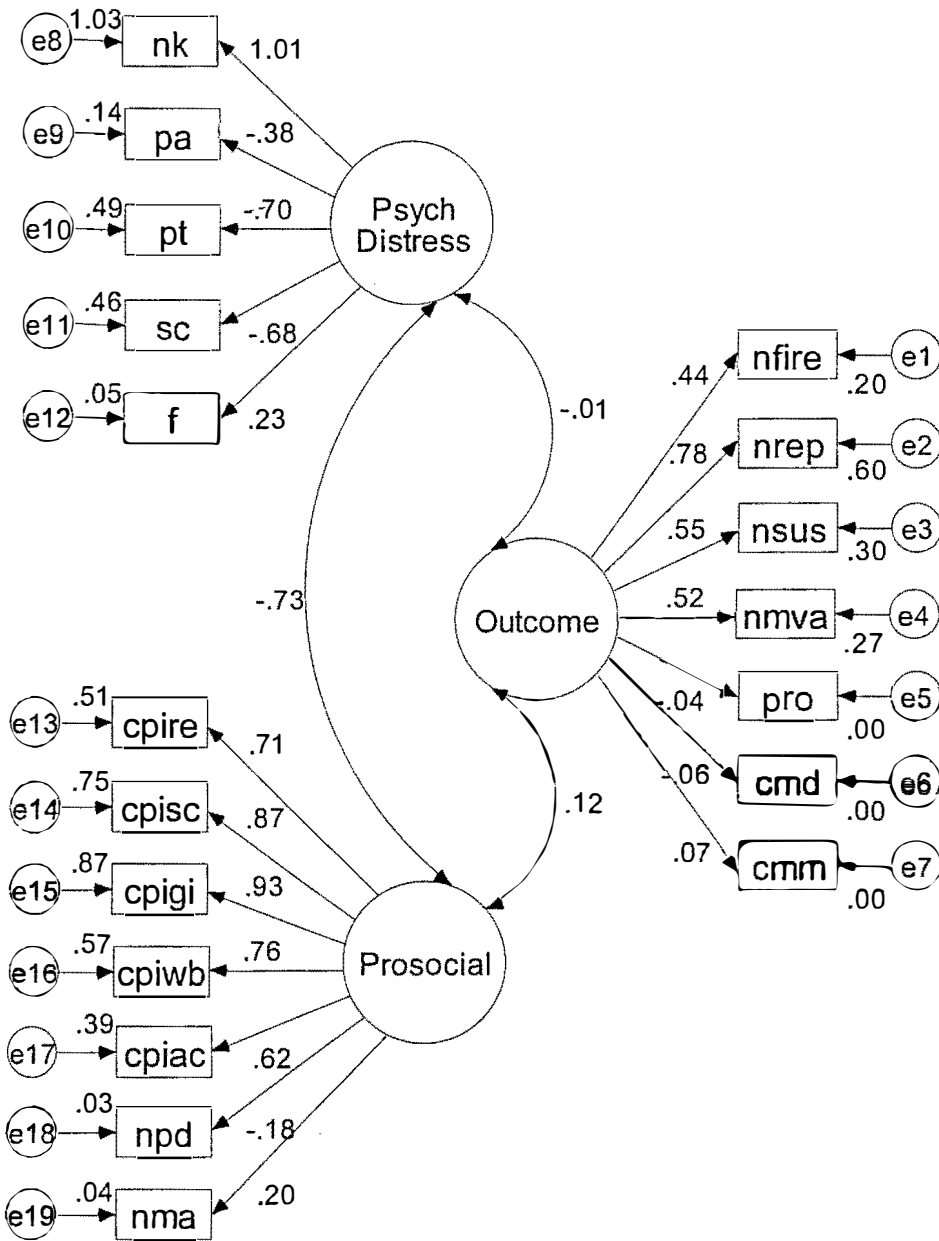


Figure 1. Initial measurement model using K-corrected T scores (N=218).

Also consistent with earlier concerns, K and F of the MMPI-2 may have contributed to a large negative correlation between the prosocial and psychological distress latent factors. For that reason and because the parameter estimate of F was low and the error variance of K was out of range, both variables were eliminated from the model.

All positive outcome variables (i.e. CMD, CMM, and PRO) had parameter estimates with absolute values less than .10. Parameter estimates for the observed variables associated with the outcome construct appeared to suggest that the negative variables and positive variables were best represented by separate latent constructs. The next model, therefore, included separate negative and positive outcome latent variables.

With these modifications, the revised measurement model, as shown in Figure 2, was tested. Path loadings fell within an acceptable range and fit measures indicated that the model fit the data well ( $\chi^2 = 155.26$ ,  $df = 98$ ;  $\chi^2/df = 1.58$ ; RMSEA = .052; CFI = .953; TLI = .942). There were moderate to high factor loadings linking the negative outcome latent variable to its associated observed variables. Parameter estimates ranged from .44 for firings to .79 for reprimands. Moderate factor loadings linked the positive outcome latent variable to its associated observed variables. Parameter estimates ranged from .43 for

major commendations to .57 for promotions. Moderate to high factor loadings linked the psychological distress latent variable to its associated observed variables. Parameter estimates ranged from .52 for the Paranoia subscale to .89 for the Schizophrenia and Psychasthenia subscales. Finally, moderate to high factor loadings linked the prosocial latent variable its associated observed variables. Parameter estimates ranged from .62 for the Achievement via Conformity subscale to .95 for the Good Impression subscale.

The negative outcome latent variable had a small positive relationship (parameter estimate  $< .10$ ) with both the positive outcome latent variable and the psychological distress latent variable. Small negative relationships (parameter estimates  $\leq -.10$ ) were found linking the negative outcome latent variable with the prosocial latent variable, as well as the positive outcome latent variable with the psychological distress latent variable. There was a moderate negative relationship (parameter estimate of  $-.33$ ) between the positive outcome latent variable and the prosocial latent variable. Finally, a moderate positive relationship (parameter estimate of  $.33$ ) was found between the psychological distress latent variable and the prosocial latent variable (See Figure 2).

*Structural Model.* The structural model, as shown in Figure 3, was evaluated to assess the direct effect of the psychological distress latent construct

and the prosocial characteristics latent construct on the positive and negative outcome latent constructs. Results indicated a good fit of the data to the model ( $\chi^2 = 155.27$ ,  $df = 99$ ;  $\chi^2/df = 1.56$ ; RMSEA = .051; CFI = .954; TLI = .944).

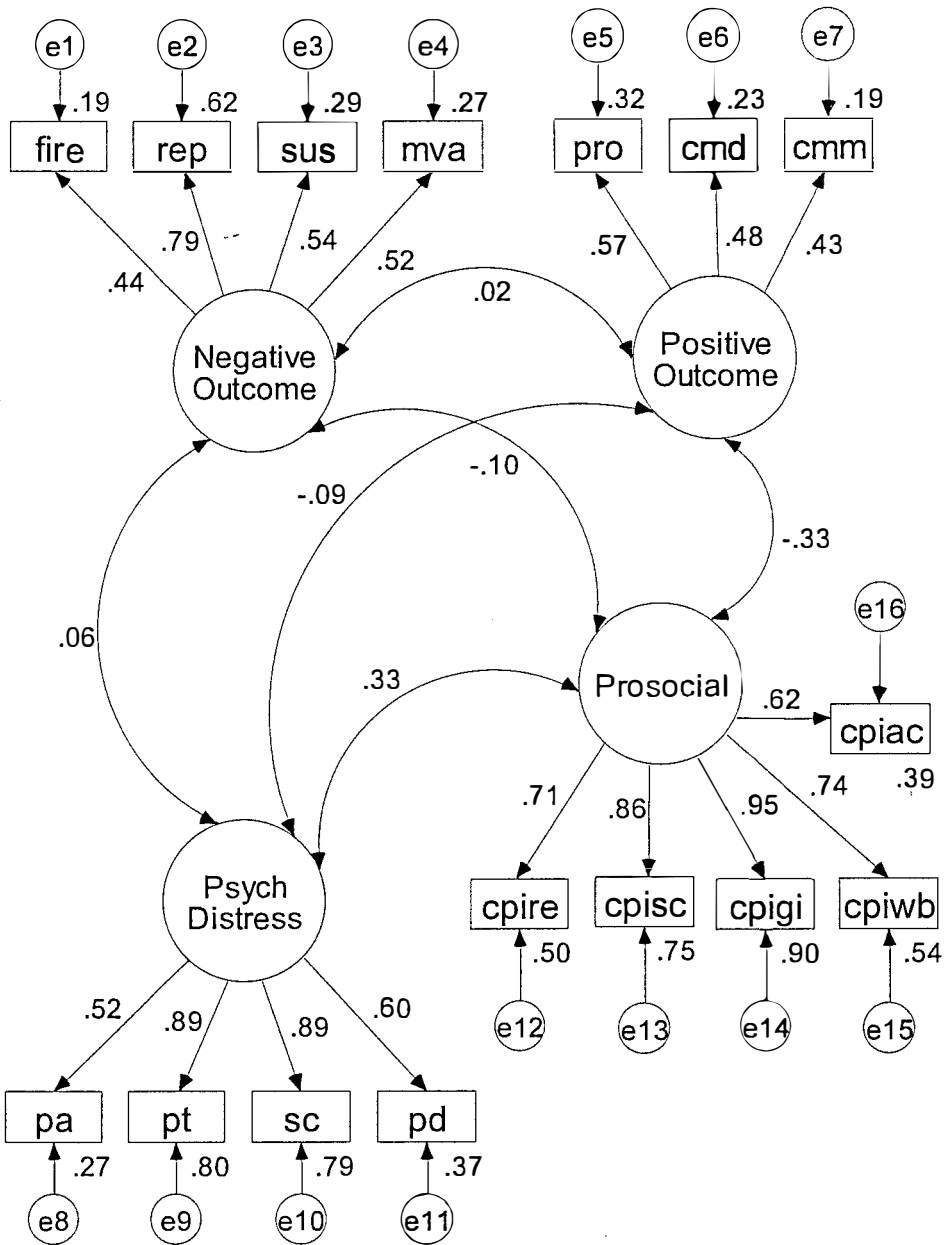


Figure 2. Modified measurement model using K-corrected T scores (N=218).

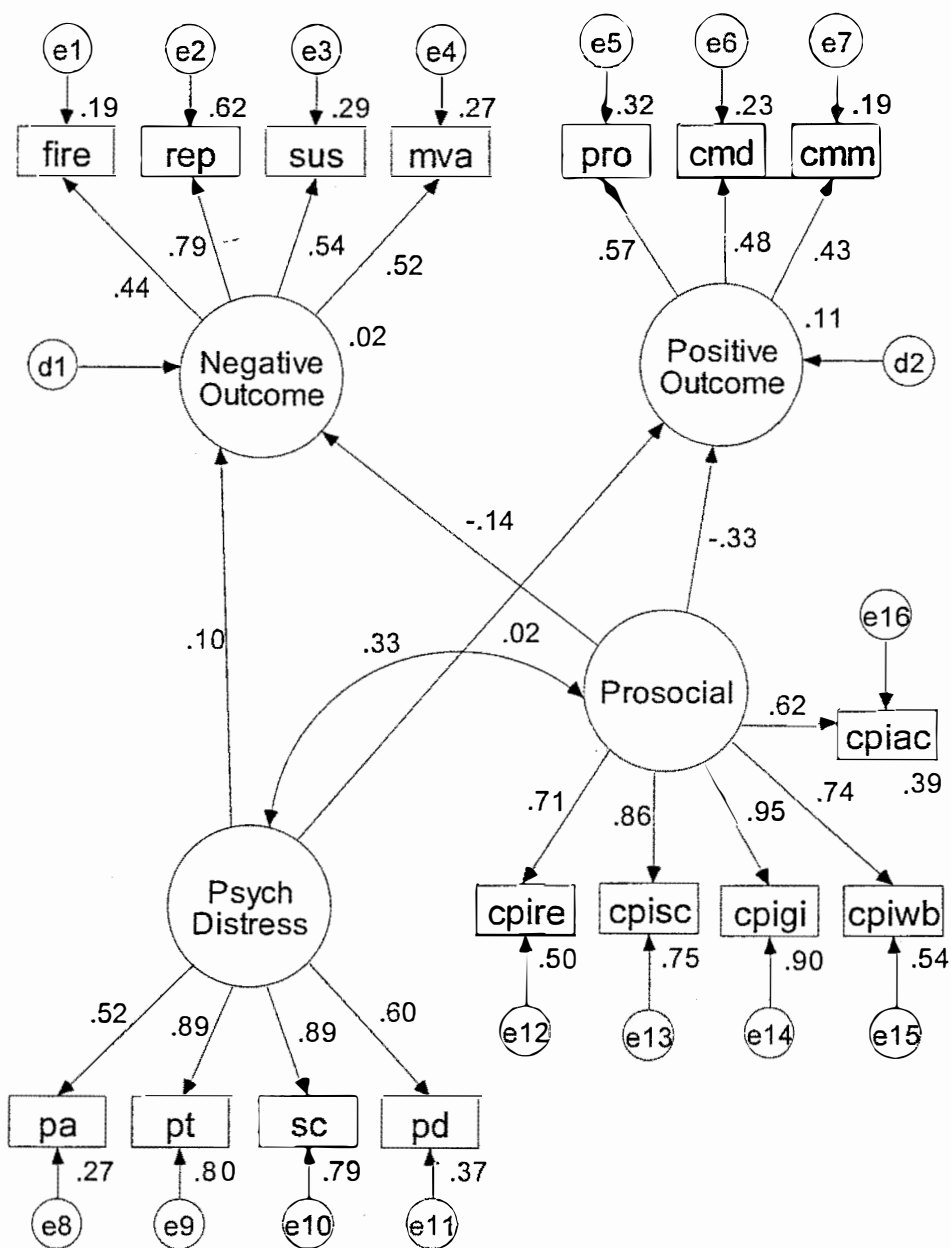


Figure 3. Structural model using K-corrected T scores (N=218). Testing the effect of psychological distress and prosocial characteristics on positive and negative outcome.

Results seen in the model in Figure 3 do not provide strong support for the hypothesis that the psychological distress construct is correlated with instances of positive and negative outcome. Though the relationship between psychological distress and negative outcome is in the expected direction, it is small and not statistically significant at .10. The path between psychological distress and positive outcome shows no relationship with an estimate of 0.02. Hypotheses regarding the prosocial construct are even less well supported because the relationship between the prosocial construct and positive outcome is moderate in size, but in the unexpected direction at -.33. Although the relationship between the prosocial construct and negative performance is in the expected direction, it is small and not significant at -.14.

*Follow-up Analyses.* Further analyses were conducted in an attempt to better understand the unexpected findings. It was thought that officers with fewer than two years of experience may not have had sufficient opportunity to have at least one positive or negative outcome event. Data were, therefore, analyzed eliminating officers with less than two years of experience. This analysis revealed no significant differences in terms of the factor loadings between observed and latent variables, or in terms of the relationship between latent variables. Analysis was also conducted with the female subjects deleted, but the results were



unchanged. Even if women legitimately do represent a different sample, their low frequency in the current study was not likely to have a strong impact on the results. In addition, the model was run with minority subjects eliminated with no significant alteration in the results.

The unexpected moderate negative correlation between the prosocial latent construct and positive outcome prompted closer inspection of the CPI scores that are represented in the prosocial construct. Several participants had scores in the extreme high range on CPI scales Re, Sc, and Gi. Using direction from the CPI interpretive guide (McAllister, 1996), Re, Sc, and Gi were transformed in an attempt to correct for possible curvilinear relationships with their latent construct and ultimately with the outcome variables. For each of these three scales, extreme high scores were recoded as scores in the low range. The conversion resulted in significantly reducing the factor loadings between all three scales and the prosocial factor. Therefore, it was decided that the CPI scales should be left unchanged, including extreme high scores. This decision serves to maintain good model fit, but this post-hoc investigation reveals that the “prosocial” latent variable requires redefinition.

Although data transformations were conducted to minimize the impact of years of service on outcome, there was still reason to suspect that years of

experience could have a confounding impact on the results. To test for possible mediating effects of years of experience on outcome, a structural model was created including years-of-experience as an observed variable with paths connecting from the psychological distress and prosocial latent variables, and going to the positive and negative outcome variables. Results indicated a good fit of the data to the model ( $\chi^2 = 182.82$ ,  $df = 111$ ;  $\chi^2/df = 1.65$ ;  $RMSEA = .052$ ;  $CFI = .947$ ;  $TLI = .935$ ) (See Figure 4). The years-of-experience latent construct has a high positive relationship (parameter estimate of .77) with the positive outcome latent construct and a moderate to high negative relationship (parameter estimate of -.51) with the prosocial latent construct. A small negative relationship (parameter estimates  $\leq -.10$ ) was found between the years-of-experience latent construct and both the negative outcome and psychological distress latent constructs. Compared to the structural model in Figure 3, there is a significantly lower parameter estimate between the prosocial construct and the positive outcome construct when years-of-experience is included as a mediating variable (from -.33 in Figure 3 to .03 in Figure 4). There were no other significant changes in observed variable-latent variable relationships or the relationships between latent constructs.

One final structural model was evaluated in order to simplify and clarify the nature of the observed variable - latent variable relationships as well as to more clearly illustrate the primary findings in terms of the relationships between latent variables and years-of-experience. Results indicate a good fit of the data to the model ( $\chi^2 = 108.23$ ,  $df = 72$ ;  $\chi^2/df = 1.50$ ; RMSEA = .048; CFI = .961; TLI = .951) (See Figure 5). Parameter estimates among the latent variables and years-of-experience remain largely unchanged from those in Figure 4 despite the elimination of Re, Wb, and Ac from the model, and the elimination of paths connecting the psychological distress latent variable to years-of-experience, the negative outcome latent variable, and the positive outcome latent variable. The rationale for renaming the “prosocial” latent variable is elucidated in the discussed.

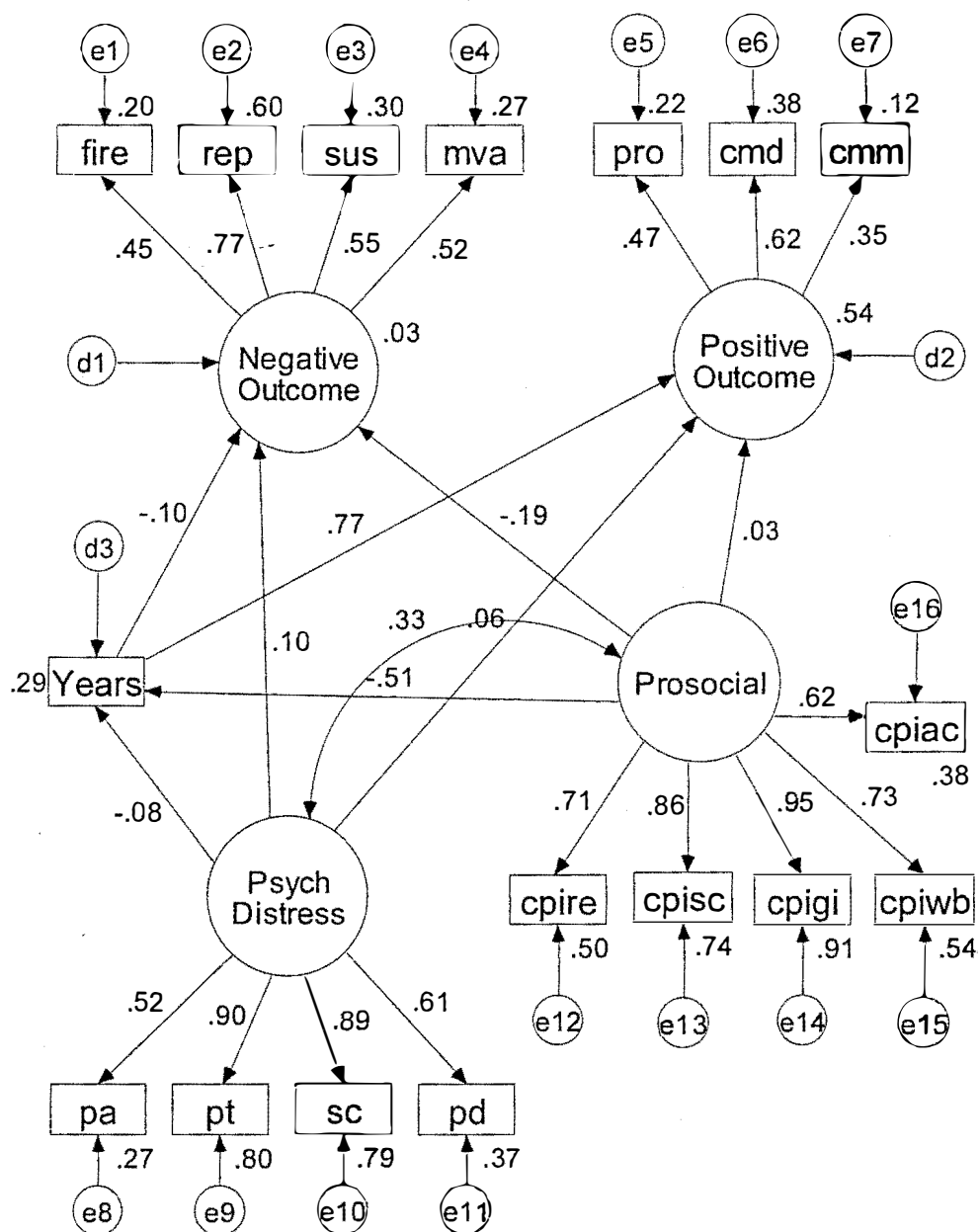


Figure 4. Structural model with years-of-experience as a mediator variable using K-corrected T scores (N=218). Testing the effect of psychological distress and prosocial characteristics on positive and negative outcome.

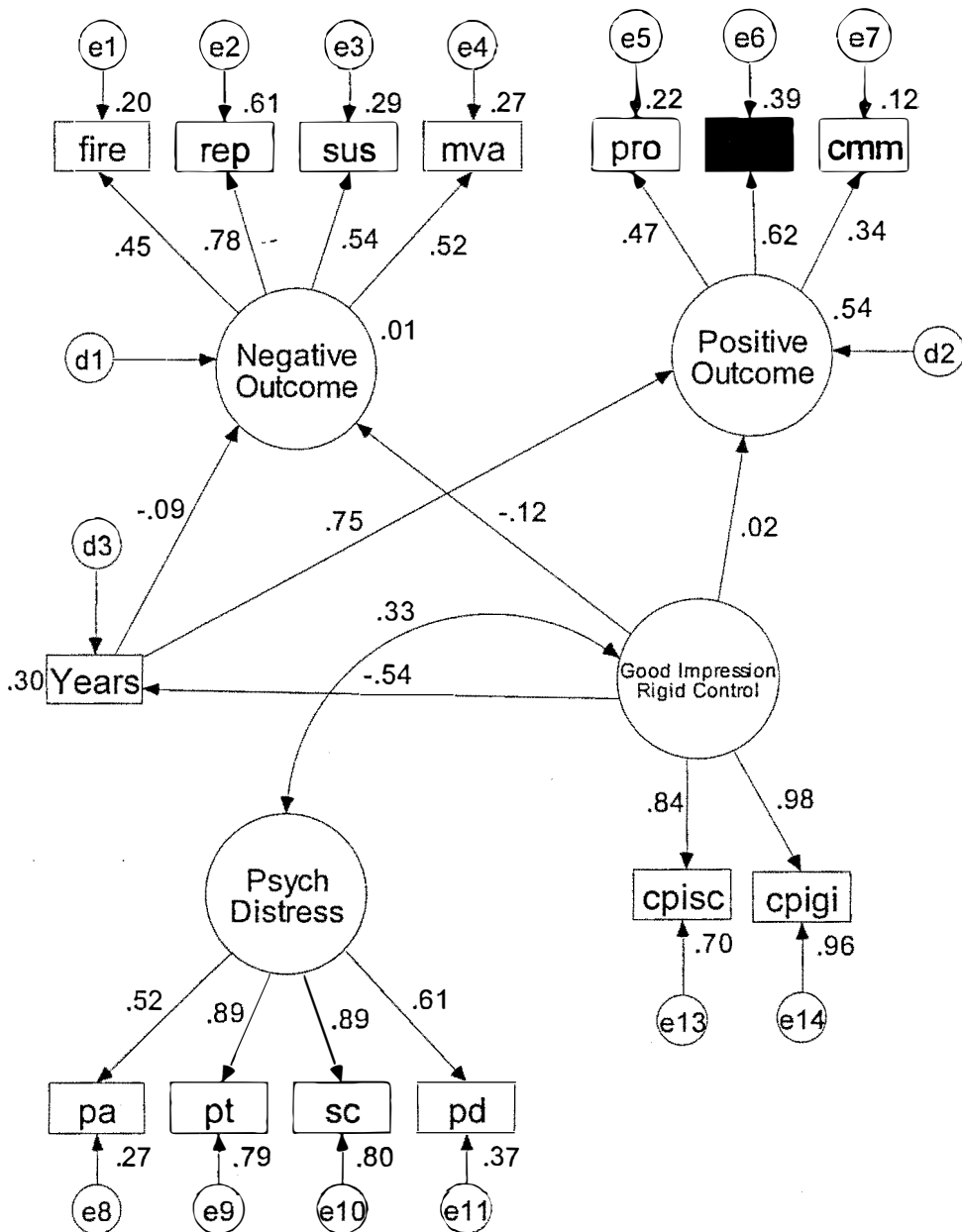


Figure 5. Simplified structural model with years-of-experience as a mediator variable using K-corrected T scores (N=218). Testing the effect of good impression and rigid control characteristics (modified label for the latent variable) on positive and negative outcome.

*Additional Results.* The following information is provided for future researchers. Tables 4 and 5 are included to provide the intercorrelations represented in the structural equation modeling analyses included in this study. All correlations between individual predictor scales and outcome variables are included. Composite variables were also included for both positive performance indicators and negative performance indicators (See Tables 4 and 5).

	F	K	Pd	Pa	Pt	Sc	Ma
MVA	.099	.070	.039	.048	.096	.103	-.100
REP	.155*	-.031	.098	-.079	.034	.010	-.049
SUS	.063	.009	.084	.041	.024	.033	.013
C-NEG	.140*	.019	.097	.001	.066	.061	-.059
CMD	.250**	-.152*	-.067	-.119	-.115	-.083	.010
CMM	.050	.064	.059	-.025	.071	.119	-.096
PRO	.142*	-.128	-.140*	-.038	-.092	-.036	-.072
C-POS	.256**	-.148*	-.078	-.116	-.107	-.059	-.023

\*. Correlation is significant at the .05 level (2-tailed).

\*\* . Correlation is significant at the .01 level (2-tailed).

Table 4. K-corrected MMPI T-scores correlated with individual and composite outcome variables (C-NEG is a composite variable, which is comprised by MVA+REP+SUS and C-POS is a composite variable, which is comprised by CMD+CMM+PRO).

	Re	Sc	Gi	Wb	Ac
MVA	-.063	.051	.015	-.098	.044
REP	-.107	-.067	-.071	-.173**	.012
SUS	-.128	-.082	-.050	-.201**	-.049
C-NEG	-.130	-.045	-.048	-.206**	.003
CMD	-.165*	-.211**	-.239**	-.163*	-.045
CMM	.077	-.028	-.002	.015	-.010
PRO	-.072	-.170*	-.248**	-.086	-.003
C-POS	-.144*	-.224**	-.260**	-.157*	-.041

\*. Correlation is significant at the .05 level (2-tailed).

\*\* . Correlation is significant at the .01 level (2-tailed).

Table 5. CPI T-scores correlated with individual and composite outcome variables (C-NEG is a composite variable, which is comprised by MVA+REP+SUS and C-POS is a composite variable, which is comprised by CMD+CMM+PRO).

Additional analyses showed that correlations between MMPI scores and CPI scores were dramatically different depending on if the K-corrected T scores were used or the non-K-corrected T scores were used (See Table 6). Finally, correlations between MMPI scores and outcome measures were shown to be somewhat different depending of if the K-corrected T scores were used or the non-K-corrected T scores were used (See Table 7). The results seen in Tables 6 and 7 indicate that the use of K-corrections has a significant impact on research findings and in some cases changes the level of statistical significance of a relationship.

Further research investigating the impact of the use of K-corrected versus non-K-corrected MMPI-2 T scores in personnel selection is clearly needed.

CPI Scale	MMPI-2 T scores							
	Pd (4)		Pt (7)		Sc (8)		Ma (9)	
	K	NoK	K	NoK	K	NoK	K	NoK
Re	.097	-.161*	.272**	-.425**	.212**	-.449**	-.189**	-.294**
Sc	.129	-.166*	.283**	-.564**	.242**	-.547**	-.288**	-.410**
Gi	.147*	-.182**	.307**	-.606**	.266**	-.607**	-.130	-.275**
Wb	.104	-.220**	.247**	-.617**	.234**	-.577**	-.214**	-.335**
Ac	-.060	-.269**	.195**	-.384**	.110	-.454**	-.073	-.185**

Table 6. Correlations between CPI scores and MMPI-2 T scores with and without the K-correction.



Outcome Variable	MMPI-2 T scores							
	Pd (4)		Pt (7)		Sc (8)		Ma (9)	
	K	NoK	K	NoK	K	NoK	K	NoK
MVA	.039	.018	.096	.011	.103	.007	-.100	-.121
REP	.098	.124	.034	.093	.010	.048	-.049	-.032
SUS	.084	.086	.024	.019	.033	.035	.013	.015
CMD	-.067	.000	-.115	.063	-.083	.120	.010	.048
CMM	.059	.034	.071	-.038	.110	.026	-.096	-.120
PRO	-.140*	-.029	-.092	.089	-.036	.137*	-.072	-.045

\*. Correlation is significant at the .05 level (2-tailed).

Table 7. Correlations between outcome variables and MMPI-2 T scores with and without the K-correction.

## Discussion

The present study was designed to minimize or eliminate many of the failings of previous research in the area of personality assessment and police officer performance. It was also hoped that through the use of sound methodology, results would provide more definitive insight into the true relationship between MMPI-2 and CPI scores and police officer outcome. Although the former may have been accomplished with the current research, the latter was not accomplished as expected. Predicted and unexpected results are discussed in terms of the strengths and limitations of the current study and the use of measures of personality in the selection of police officer candidates.

### *Latent Constructs for Outcome*

Outcome measures were best represented by two latent constructs. Consistent with expectations, firings, suspension days, written reprimands, and motor vehicle accidents all loaded on a single latent construct representing negative police officer outcome. In the final structural model (Figure 4) written reprimands had the strongest factor loading with the negative performance latent variable accounting for 60% of its variance. It is intuitive that the same officers who have written reprimands in their record also have other instances of negative outcome such as suspensions and eventually terminations.

Also consistent with prior expectations, promotions, written commendations, and major awards all loaded on a single latent construct best described as positive police officer outcome. Parameter estimates were more modest for this latent construct with the positive outcome latent variable accounting for 22% of the variance for promotions and 38 % of the variance for written commendations. These results support the intuitive expectation that officers who receive a high number of written commendations are also being promoted and receive awards at a higher rate than other officers. The low correlation between the positive outcome latent construct and the negative outcome latent construct indicates that there is no consistent relationship between the positive performance indicators and the negative performance indicators included in the present study. The lack of relationship between positive and negative outcome in this research is likely related to overall low frequency of outcome occurrences. Utilization of more frequent and less extreme indicators of positive and negative performance would likely result in a stronger relationship between the two. For example, if sick days, injuries, and instances of informal verbal counseling were included as negative performance indicators, and good supervisor ratings were included as positive performance indicators, there would

likely be more of a relationship between the positive and negative performance constructs.

Future research could benefit from using a similar multiple indicator approach to investigate successful and unsuccessful officers. There were, however, drawbacks to exclusively using the types of indicators of performance included in the current study rather than also relying on other indicators of outcome that are more frequently occurring, such as sick days and the more subjective supervisor or peer ratings. These drawbacks will be discussed further in the section addressing the homogeneity of the sample.

#### *Latent Constructs for Personality Scales*

With some alteration to the originally proposed measurement model, two moderately correlated latent constructs were identified among the personality subscales (i.e. psychological distress and prosocial characteristics). Both the psychological distress latent construct and the prosocial latent construct underwent some alteration in terms of the ultimate combination of observed variables represented. The psychological distress factor, despite some alteration, retains the same basic descriptive characteristics as originally proposed. The scores on the MMPI scales that load on the psychological distress factor, however, were in the normal range, indicating that subjects from the current sample represent a range of

psychological distress from the low end to the moderate-normal range. The difficulty with using specific MMPI scales in personnel selection encountered in the current study was also encountered in previous research (e.g. Butcher, 1994; Saxe & Reiser, 1976). Although Butcher (1994) liberally recommended using interpretations of moderate elevations ( $T > 60$ ) in cases of personnel selection, in the current study, 90% of participants had  $T < 57$  for MMPI-2 scales Pa and Pd and  $T < 53$  for MMPI-2 scales Pt and Sc. The infrequency of scores even in the moderately elevated range contributed to the low relationship between the psychological distress factor and outcome.

The basic scales of the MMPI-2 are not the most sensitive tools for predicting performance in a largely healthy population or one motivated to look psychologically well adjusted. The MMPI-2, however, is still useful when hiring for safety sensitive positions because it may screen out individuals with significant psychopathology. It may be the case that the use of the MMPI in selecting the police officers included in the current study, in part, reduced the sensitivity of the MMPI to predict future outcome among those selected for hire. Screening out applicants with more significant psychopathology likely reduced the number of problem officers hired, thereby resulting in an overall paucity of negative outcome

indicators, and leaving only profiles with scores that rarely approach the clinical or interpretable range.

The prosocial factor was originally proposed as an antisocial/prosocial continuum based strongly on the expectation that MMPI scales Pd and Ma would load on that factor and that they represented antisocial characteristics. Given the final combination of CPI scales represented by this latent construct, a new definition and label is warranted for this construct. Follow up analyses revealed that a large number of participants had scores in the extreme high range on CPI scales Re, Sc, and Gi. Using the CPI interpretive manual as a guide, the latent construct seems to be a combination of traits such as conformity, orderliness, rigid control, rigid adherence to rules, high desire to please others, and desire to look good in the eyes of others (McAllister, 1996). The latent construct accounts for 95% of the variance in the Gi subscale and 86% of the variance of the Sc subscale. A more appropriate label for the latent variable, therefore, seems to be “desire to look good,” “rigid-control,” or a more comprehensive label such as “good impression and rigid control.” For the sake of parsimony, the “good impression and rigid control” latent construct may be more simply represented by only the Gi and Sc scales. A structural model was run eliminating Re, Wb, and Ac from the model with no significant change in model fit or any parameter estimates. This

indicates that Gi and Sc are not only the two strongest contributing scales to the latent construct, but that they alone are sufficient to define the construct.

Redefining what was previously proposed as the prosocial latent construct helps explain the moderate positive correlation between it and the psychological distress construct. Some correlation between the two proposed latent constructs was expected due to the overlap of items between the CPI and MMPI (171 items overlap between the original MMPI and current CPI and 158 items overlap between the MMPI-2 and current CPI) (Gough, 1996). Some correlation was also expected based on previous research that revealed intercorrelations between the MMPI and CPI subscales included in the present study. The moderate positive correlation of .33 between the psychological distress factor and prosocial factor, however, was unexpected because previous research revealed relatively low and negative correlations between the MMPI and CPI subscales included in the current study (Gough, 1996). Issues related to the approach participants take in selection assessments, unique characteristics of a sample of police officer candidates, and more accurately defining the two constructs all help explain this unexpected finding.

Previous findings revealing low negative correlations between the MMPI and CPI subscales in question were extrapolated from a normative sample. MMPI

scores for police officers have been shown to be less elevated when compared to the normative sample (Carpenter & Raza 1987). MMPI scores for the current sample were also low and rarely approached the clinical range. Furthermore, the test taking approach in a selection environment likely impacted scores on the MMPI and CPI in ways that diverge from a normative sample. It is likely that the desire to be seen in a favorable light contributed to elevations in CPI scores for the current sample.

Given that the latent variable does not represent prosocial characteristics, it is less surprising that there was a moderate correlation between the psychological distress factor and the “prosocial” factor. Participants who attempt to portray themselves in an unrealistically favorable light and are rigidly controlled are likely to be somewhat elevated in terms of psychological distress. It should be noted, however, that there can be problems associated with over-interpreting normal range scores on MMPI scales and that the psychological distress factor does not represent a high level of psychological distress for the participants in the current study. Given this limitation in the current study, the general hypothesis that high psychological distress in police officer candidates leads to negative outcome, and that it results in less frequent positive outcome still may be tenable.



The unexpected result of a relatively high negative relationship between the proposed prosocial factor and positive outcome is also less surprising when the prosocial factor is relabeled as the good impression and rigid control factor. Follow up analyses indicate that as tendencies toward making a good impression and rigid control increase, the likelihood of commendations and promotions decreases. Officers with these characteristics, or this style of responding to CPI items, were not eliminated by the hiring process, and they did not have an increased likelihood of extreme negative performance indicators, as indicated by the low negative correlation with the negative outcome latent construct. Despite successful avoidance of disciplinary action, however, officers with these traits do not tend to go on to be the most successful officers. A tendency toward rigid rule following would likely assist an officer in the avoidance of disciplinary action, while too much rigidity and excessive focus on making a good impression is not consistent with the leadership qualities required to achieve promotion or to be seen as a top performer.

Since follow up investigation revealed that the current study did not effectively identify a prosocial characteristics factor, the hypothesis that prosocial traits would predict positive outcome was not adequately tested in the present

study. Future research utilizing a more valid indicator of prosocial traits should reveal a connection between prosocial characteristics and positive outcome.

### *Confound of Years of Experience*

Preliminary data transformations were conducted to minimize the impact of years of service on outcome and to normalize the distribution of outcome scores. All outcome variables were divided by years of experience for each subject and transformed using the square root function, but there is still some reason to suspect that years of experience could have a confounding impact on the results. In order to test for possible mediating effects of years of experience on outcome, a structural model was tested using years-of-experience as an observed variable. Despite the data transformations to adjust for years-of-experience, years-of-experience was still found to have a significant impact on outcome.

Including years-of-experience in the structural model significantly reduced the parameter estimate between the prosocial (more appropriately named “good impression and rigid control”) construct and the positive outcome construct, while revealing strong direct relationships with both. Given the design of the current study it is difficult to determine, with certainty, the reasons for the strong positive relationship between years-of-experience and positive outcome, and the strong

negative relationship between years-of-experience and the good impression and rigid control latent variable.

The data transformation used in the current study to reduce the effect of years of experience on outcome (outcome variable frequency / years of experience) assumes that each year with the department is equal in terms of opportunity to achieve outcome indicators. Results in Figure 4 indicate that this is not true for positive outcome indicators. There seems to be a relationship between positive outcome indicators and years-of-experience in which it is more likely that an officer would receive a promotion or commendation in the fifth, sixth, and seventh year on the force as compared to the first, second, or third. This finding, when added to the finding that years-of-experience has a strong negative relationship with the good impression and rigid control latent construct suggests two possible interpretations. This effect may reflect test-taking differences between cohorts or it may reflect attrition of officers who are rigidly controlled and who have very high desire to make a good impression.

Although years-of-experience in the current study is not strictly an indicator of whether psychological testing was done more recently or longer ago, there is likely some relationship between the two. This effect could represent a cohort effect in which officers who were evaluated longer ago generally have less

elevated scores on CPI scales and that the more recent test takers have more elevated CPI scores. Higher scores on these scales, therefore, could indicate that the more recent group of applicants showed an increased motivation to attempt to answer questions in the perceived favorable direction. Although not probable, this is a plausible explanation if in the two departments included in this research, the influence of the psychological testing process, and the applicant's awareness of this influence, has increased over the years covered in the study. Clinical observation and previous research, however, suggest that police officer candidates and applicants for other positions alike have always tended to respond in the perceived favorable direction.

Although it could be argued that the effect of years-of-experience in the measurement model simply reflects that more seasoned officers are more frequently rewarded with indicators of positive outcome, longevity in the department can be interpreted as a positive indicator of officer performance in its own right. These results may indicate that officers who are rigidly controlled and focused too strongly on making a good impression do not tend to achieve positive outcome indicators and that they have abbreviated careers with one department. Although these officers did not seem to have a high rate of firings, they may have been more likely to leave the department due to lack of progression. Officers who

are rigidly controlled and highly focused on making a good impression may not express the autonomous behavior and independent thinking required to achieve high praise and promotion. Someone with these traits who expends significant energy attempting to impress others and succeed in the department may feel overlooked when watching his peers receive accolades and promotions, which may lead to a decision to leave the department.

Future research may help to clarify these issues. Research investigating the effect of extreme high scores on CPI scales Gi and Sc on positive police officer outcome with other police officer populations would give information regarding the level of generalizability of the current findings. Moreover, research specifically designed to utilize attrition rates as a measure of outcome could serve to clarify and expound upon the current results.

### *Objective versus Subjective Measures of Performance*

As discussed earlier, drawbacks to using more subjective measures of outcome such as supervisor ratings include suspect reliability and validity. The outcome measures referred to as less subjective and more face valid used in the current study may have similar problems with reliability. Although these outcome variables are represented by discrete occurrences or events and have high face validity, they are still based in some respect on subjective appraisals of

performance. One of the benefits of using more overtly subjective appraisals is that they can provide greater variability. The outcome measures used in the present study are valuable due to their direct face validity, but the extreme infrequency of occurrences can be a limiting factor in terms of identifying personality scales as predictors of outcome.

In addition to the measures of performance used in the current study, future research would benefit from incorporating other measures of performance that range in level of subjectivity. Ideally these performance indicators would be selected or generated such that they limit halo effects, problems of unreliability, and demand characteristics. One possibility would be to generate an outcome indicator by having selected participants anonymously evaluated by supervisors, peers, and subordinates. Citizens, with whom the officers interact, could also be asked to complete a questionnaire regarding their satisfaction and opinions. Researchers could also rate videos of traffic stops based on selected criteria of skillful execution. The unfortunate reality of this type of research is that obtaining ideally reliable and valid performance indicators is often prohibitively expensive in terms of time and resources.

One feasible way to significantly improve the quality of police officer selection research may be for departments to begin keeping shadow records of

officer performance that do not impact promotion or disciplinary action. These records would be maintained solely for research purposes. Since officer personnel files are currently used for promotion and discipline, there are many potentially useful pieces of information that are not maintained in these records, or they are eliminated after a period of time. Informal instances of verbal correction, for example, never appear in a personnel file. Even formal recorded instances of verbal counseling and some written reprimands are eliminated from the personnel file after a period of service with no further negative action. If these and other instances of more routine positive and negative performance were recorded and maintained for future research, the predictive ability of existing police officer selection procedures would likely improve.

Shadow records could also be kept up to date regarding issues that may not lead to disciplinary action, but are of interest in terms of economically managing a police department. Frequency of injuries, sick days, and damage to department property are examples of such outcome variables. Although this would not be an ideal solution to the problem of finding good outcome measures, it would serve to vastly increase the variability of indicators already utilized and seen as face valid. These types of outcome indicators could increase the sensitivity of MMPI, CPI, or

other pre-hire screening tools for the prediction of positive and negative performance.

### *Homogeneity of Sample*

Although the two police departments included in the current study were selected based on their similarities and both were included to obtain an adequate sample size, there may be some problems associated with including participants from different departments in the same research. Participants from the two departments were found to be similar in terms of a variety of demographic factors, and their responses to the MMPI and CPI, but some observable differences remained. Preliminary analyses revealed that there were some differences in terms of years of experience, age between participants, and frequency of some outcome indicators.

The effect of the mean difference of 2.29 in terms of years of experience was likely minimized due to the data transformation applied to reduce the impact of years of service. The mean difference in age of 4.53 years between the two subsamples may be evidence that the two samples are not representative of the same population. Statistically significant differences between the two departments were found for frequency of commendations, major awards, and reprimands. Although these differences were statistically significant, it could be argued that the small



actual difference in mean scores of less than three tenths of a single outcome event indicates low practical significance. Nevertheless, one explanation for these differences may lie in the level of subjectivity associated with the outcome measures. The effect of subjectivity in measures can be seen at an individual level, as in the example of individual supervisor ratings, or it can be seen at a more organizational level. Different institutionalized standards may be applied across departments when determining negative and positive consequences. For example, misuse of a departmental issue cellular phone may result in a suspension in one department and a written reprimand or verbal counseling in another department.

Results from the current study support previous findings that suggest there may be problems associated with generalizing results to diverse departments. Previous research identified differences associated with size of department, type of community served, and the type of law enforcement activities performed (Bartol, 1982; Carpenter & Raza, 1987; Hargrave & Hiatt, 1987; Hargrave, Hiatt, & Gaffney, 1986; Saccuzzo, et al., 1974; Saxe & Reiser, 1976), which were controlled in the current study. Knowledge of these interdepartmental differences should be used to inform future research, as well as to caution clinicians regarding the generalizability of specific findings to practice. This type of research would be

most effective if departments of sufficient size and with adequate resources commissioned research to be done locally and to be applied to future hiring.

### *Conclusion*

Despite its potential limitations, the MMPI-2 still appears to be a valuable tool to include in a police officer selection assessment battery. There is not yet enough evidence in the research, or in the current study, to support the use of selected sub-scales to make definitive decisions regarding fitness for duty. It is recommended that the MMPI-2 profile as a whole be considered by an experienced professional in conjunction with clinical interview and other sources of information. Given its intended use with patients with serious mental health conditions, it is also recommended that the MMPI-2 be augmented with other measures of personality intended for use with higher functioning participants. The MMPI-2 restructured (RC) scales may prove promising predictors due to their improved discriminant validity over the original MMPI-2 basic scales. Some results in the current study suggest that when the MMPI-2 basic scales are used in personnel selection, they should be interpreted without the application of the K-correction. The current study provided support for using the CPI in a pre-hire assessment battery, which adds to the previous research that has pointed to the CPI as a good option to be used in conjunction with the MMPI-2 (Borstman, 1977;

Hogan, 1971; Mills & Bohannon, 1980). There is also research to support the use of the IPI in police officer selection (Mufson & Mufson, 1998; Inwald & Susman, 1984).

The present study revealed possible latent constructs related to personality characteristics and outcome variables, but failed to show the hypothesized relationships between those constructs. Although these findings do not add to the existing literature as predicted, the research design choices made in the current study regarding sample size, criterion measures, predictive methodology, and use of other personality measures rather than the MMPI alone, can provide direction for future research investigating the true relationship between personality measurement and police officer outcome. Some results that were not predicted may also provide direction for future researchers.

The results of the current study provide insight into the relationship between officers who display rigid control and excessive focus on making a good impression, and some outcome indicators. Results indicate that although officers who display these characteristics successfully avoid disciplinary action, they do not tend to achieve commendations or promotions. There was also some evidence to indicate that officers with these traits may have higher turnover than officers with test scores that do not indicate the presence of these traits. If this specific finding

is replicated and further clarified, the CPI scales Gi and Sc may prove to be a valuable tool in the very difficult task of detecting marginally performing future officers and applicants who tend not to have long careers with one department. Successful reduction of high rates of officer turnover would lead to economic savings in terms of the hiring process, training of new officers, and the potentially costly effect of having too few experienced officers serving our communities.

References

Anastasi, A. (1988). *Psychological Testing, 6<sup>th</sup> ed.* New York: Macmillian.

Anderson, J. C., & Gerbing, D. W. (1988). Structural Equation Modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*, 411-423.

Arbuckle, J. L. & Wothke, W. (1999). *Amos 4.0 User's Guide.* Chicago IL: SmallWaters.

Barthlow, D. L., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., McNulty, J. L. (2002). The appropriateness of the MMPI-2 K Correction. *Assessment, 9*, 219-229.

Bartol, C. R. (1982). Psychological characteristics of small-town police officers. *Journal of Police Science & Administration, 10*, 58-63.

Bartol, C. R. (1991). Predictive validation of the MMPI for small-town police officers who fail. *Professional Psychology - Research & Practice, 22*, 127-132.

Bernstein, I. H., Schoenfeld, L. S. Costello, R. M. (1982). Truncated component regression, multicollinearity and the MMPI's use in a police officer selection setting. *Multivariate Behavioral Research, 17*, 99-116.

- Beutler, L. E., et al. (1985). Parameters in the prediction of police officer performance. *Professional Psychology - Research & Practice, 16*, 324-335.
- Beutler, L. E., Nussbaum, P. D., Meredith, K. E. (1988). Changing personality patterns of police officers. *Professional Psychology - Research & Practice, 19*, 503-507.
- Borstman, P. L. (1977). Assessing the California Psychological Inventory for predicting police performance. *Dissertation Abstracts International, June 37*(12-B), 6387.
- Burbeck, E. & Furnham, A. (1985). Police officer selection: A critical review of the literature. *Journal of Police Science & Administration, 13*, 58-69.
- Burkhart, B. R. (1980). Conceptual issues in the development of police selection procedures. *Professional Psychology, 11*, 121-129.
- Butcher, J. N. (1994). Psychological assessment of airline pilot applicants with the MMPI-2. *Journal of Personality Assessment, 62*, 1, 31-44
- Butcher, J. W., and Tellegen, A. (1978). Common methodological problems in MMPI research. *Journal of Consulting and Clinical Psychology, 46*, 620-628.

Carpenter, B. N., Raza, S. M. (1987). Personality characteristics of police applicants: Comparisons across subgroups and with other populations. *Journal of Police Science & Administration*, 15, 10-17.

Casico, W. F. & Valenzi, E. R., (1978). Behaviorally anchored rating scales: Effects of education and job experience of raters and ratees. *Journal of Applied Psychology*, 62, 278-282.

Cimbura, J. A. (2000). An exploration of stress in police officers: A study of the predictive value of pre-employment psychological measures in the development of stress reactions in a sample of Ontario police officers. *Dissertation Abstracts International: Section B: the Sciences & Engineering*, 60(8-B), 4281.

Cohen, R. J., Swerdlik, M. E., Smith, D. K. (1992). *Psychological testing and assessment: second edition*. Mountain View: Mayfield Publishing.

Cortina, J. M., Doherty, M. L., Schmitt, N., Kaufman, Gary, et al. (1992). The "Big Five" personality factors in the IPI and MMPI: Predictors of police performance. *Personnel Psychology*, 45, 119-140.

Costello, R. M., Schneider, S. L., Schoenfeld, L. S. (1996). Validation of a preemployment MMPI index correlated with disciplinary suspension days of police officers. *Psychology Crime & Law*, 2, 299-306.

Costello, R. M., Schoenfeld, L. S., Kobos, J. (1982). Police applicant screening: An analogue study. *Journal of Clinical Psychology, 38*, 216-221.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *Catalog of Selected Documents in Psychology, 10*, 85.

Detrick, P., Chibnall, J. T., & Rosso, M. (2001). Minnesota multiphasic personality inventory - 2 in police officer selection: Normative data and relation to the Inwald Personality Inventory. *Professional Psychology: Research and Practice, 32*, 484-490.

FitzGerald, P. R. (1987). The prediction of police performance using the MMPI and CPI. *Dissertation Abstracts International, 47(8-B)*, 3519.

Gergen, K. J., Gergen, M. M., & Meter, K. (1972). Individual orientation to prosocial behavior. *Journal of Social Issues, 8*, 105-130.

Gough, H. G., & Bradley, P. (1996). *CPI Manual Third Edition*, Palo Alto: Consulting Psychologists Press.

Greene R., (1991). *The MMPI-2/MMPI: An interpretive manual*. Boston: Allyn and Bacon.

Hargrave, G. E. & Hiatt, D., (1987). Law enforcement selection with the interview, MMPI, and CPI: A study of reliability and validity. *Journal of Police Science and Administration, 15*, 110-116.



Hargrave, G. E., Hiatt, D., & Gaffney, T. W. (1988). F+4+9+Cn: An MMPI measure of aggression in law enforcement officers and applicants. *Journal of Police Science and Administration*, 16, 268-273.

Hargrave, G. E., Hiatt, D., & Gaffney, T. W. (1986). A comparison of MMPI and CPI profiles for traffic officers and deputy sheriffs. *Journal of Police Science and Administration*, 14, 250-258.

Hathaway, S. R., & McKinley, J. C., (1989). *Minnesota Multiphasic Personality Inventory - 2*. Minneapolis: The University of Minnesota Press.

Hayduk, L. (1987). *Structural equation modeling with LISREL*. Baltimore: Johns Hopkins University Press.

Hiatt, D. & Hargrave, G. E., (1988). MMPI profiles of problem peace officers. *Journal of Personality Assessment*, 52, 722-731.

Hickman, M. J. & Reaves, B. A., (2001, May). Local police departments 1999. *Law Enforcement Management and Administrative Statistics* [on-line], Available FTP: Hostname: virlib.ncjrs.org Directory: lawenforcement.asp File: 2001 NCJ 186478.

*Hild v. Bruner*, 496 F. Supp. 93 (D. New Jersey 1980).

Hogan, R., (1971). Personality characteristics of highly rated policemen. *Personnel Psychology*, 24, 679-686.

Hogan, R. & Kurtines, W., (1975). Personological correlates of police effectiveness. *The Journal of Psychology*, 91, 289-295.

Inwald, R. E. & Shusman, E. J. (1984). The IPI and MMPI as predictors of academy performance for police recruits. *Journal of Police Science & Administration*, 12, 1-11.

Joreskog, K. G., & Sorbom, D. (1989). *LISREL 7: A guide to the program and applications*. Chicago: SPSS Inc.

Kelley, P. L., Jacobs, R. R., & Farr, J. L. (1994). Effects of multiple administrations of the MMPI for employee screening. *Personnel Psychology*, 47, 575-591.

Kornfeld, A. D. (1995). Police officer candidate MMPI-2 performance: Gender, ethnic, and normative factors. *Journal of Clinical Psychology*, 51, 536-540.

Kunce, J. T., and Anderson, W. P. (1984) Perspectives on uses of the MMPI in nonpsychiatric settings. In P. McReynolds, and G. J. Chelune (Eds.), *Advances in Psychological Assessment*, vol. 6 (pp. 41-76). San Francisco: Jossey-Bass Publishers

McAllister, L. W. (1996). *A practical guide to CPI interpretation: third edition*. Palo Alto: Consulting Psychologists Press.

*McKenna v. Fargo*, 451 F Supp. 1355 (D. New Jersey 1978).

Meier, R. D., Farmer, R. E., & Maxwell, D., (1987). Psychological screening of police candidates: Current perspectives. *Journal of Police Science and Administration*, 15, 210-215.

Merian, E. M., Stefan, D., Schoenfeld, L.S., Kobos, J. C. (1980) Screening of police applicants: A 5-item MMPI research index. *Psychological Reports*, 47, 155-158.

Mills, C. J. & Bohannon, W. E., (1980). Personality characteristics of effective state police officers. *Journal of Applied Psychology*, 65, 680-684.

Mitchell, J. V. & Pierce-Jones, J. (1960). A factor analysis of Gough's California Psychological Inventory. *Journal of Consulting Psychology*, 24, 453-456.

Mufson, D. W. & Mufson, M. A., (1998). Predicting police officer performance using the Inwald Personalty Inventory: An illustration from Appalachia. *Professional Psychology: Research and Practice*, 29, 59-62.

Muha, T. M. & May, J. R. (1973). An employment index for identifying unfit job applicants. *Journal of Community Psychology*, 1, 4, 362-365.

Nichols, R. C. & Schnell, R. R. (1963). Factor scales for the California Psychological Inventory. *Journal of Consulting Psychology*, 27, 228-235.

Pallone, N. J. (1992). The MMPI in police officer selection: Legal constraints, case law, empirical data. *Journal of Offender Rehabilitation, 17*, 171-188.

Penner, L. A., Fritzsich, B. A., Craiger, J. P., Freifeld, T. S. (1995). Measuring the prosocial personality. In J. N. Butcher and C. D. Spielberger (Eds.), *Advances in personality assessment, vol. 10* (pp. 147-163). San Francisco: Jossey-Bass Publishers

Pugh, G., (1985). The California Psychological Inventory and police selection. *Journal of Police Science and Administration, 13*, 172-176.

Riddle, R. E. (2000). The relationship between burnout, death anxiety and aggression in police officers. *Dissertation Abstracts International: Section B: the Sciences & Engineering, 60(9-B)*, 4906.

Rushton, J. P., Chrisjohnm R. D., & Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Personality and Individual Differences, 2*, 293-302.

Saccuzzo, D. P., Higgins, G., Lewandowski, D. (1974). Program for psychological assessment of law enforcement officers: Initial evaluation. *Psychological Reports, 35*, 651-654.

Saxe, S. J., Reiser, M. (1976). A comparison of three police applicant groups using the MMPI. *Journal of Police Science & Administration*, 4, 419-425.

Scogin, F. & Beutler, L. E. (1986). Psychological screening of law enforcement candidates. In P. A. Keller & L. G. Ritt (Eds.), *Innovations in clinical practice* (Vol, 5, pp 317-330). Sarasota, FL: Professional Resources Exchange.

Thomas-Riddle, F. R. (2000). The relationship between life stress, work stress, and traumatic stress and burnout and cynicism in police officers. *Dissertation Abstracts International: Section B: the Sciences & Engineering*, 60(9-B), 4914.

Weed, N. C. & Butcher, J. N. (1992). The MMPI-2: Development and research issues. In J. C Rosen & P McReynolds (Eds.), *Advances in Psychological Assessment* (Vol, 8, pp 131-163). New York: Plenum Press.

Wasson Dralle, P. & Baybrook, R. M. (1985). Screening of police applicants: A replication of a 5-item MMPI research index validity study. *Psychological Reports*, 57, 1031-1034.

Wilikofsky, A. S. (1986). MMPI based employment decisions for high-risk positions in nuclear power settings. *Dissertation Abstracts International: Section B: the Sciences & Engineering*, 48(2-B), 576.