THE NEO PI-R, INWALD PERSONALITY INVENTORY, AND MMPI-2 IN THE PREDICTION OF POLICE ACADEMY PERFORMANCE: A Case for Incremental Validity†

JOHN T. CHIBNALL††
Saint Louis University

PAUL DETRICK
Florissant Psychological Services

ABSTRACT: The utility of personality inventories for predicting successful police officer performance is an important issue in the law enforcement employment selection process. The purpose of the present study is to evaluate the incremental validity of the Minnesota Multiphasic Personality Inventory (MMPI-2), the Inwald Personality Inventory (IPI), and the Revised NEO Personality Inventory (NEO PI-R) for predicting police academy performance of 79 recruits. Results indicate that each inventory contributes significantly to prediction of academic performance in the academy, while only the NEO PI-R predicts physical performance. The addition of the NEO PI-R to selection processes that already incorporate the MMPI-2 and IPI may be useful for enhancing prediction of police officer performance.

INTRODUCTION

The use of psychological and personality inventories to screen and select police officer candidates has a long history (Inwald, 1987). The primary motivation for using these instruments is to limit liability. That is, psychological screening eliminates job applicants who are unfit for police work or who, by dint of a psychopathological condition, are considered high risks for behavioral, psychological, or disciplinary problems. In addition to identifying negative characteristics that may interfere with police work, personality inventories may predict success by identifying positive profiles that are associated with exemplary po-
lice officer performance. This issue of the relative utility of "screening out" high-risk candidates versus "screening in" potentially excellent applicants continues to warrant attention in police officer selection (Metchik, 1999).

The role of personality assessment in this debate is significant. The best inventory or combination of inventories would identify both high-risk and potentially excellent candidates with an acceptable level of predictive validity. The latter use, however, has garnered less attention than the former. Two inventories that have received much attention with respect to screening and prediction in law enforcement are the Minnesota Multiphasic Personality Inventory (MMPI), along with its current revision (MMPI-2), and the Inwald Personality Inventory (IPI). A third personality inventory, the Revised NEO Personality Inventory (NEO PI-R), has emerged recently. However, it is the least researched and used of the three instruments with respect to police officer selection and performance prediction.

Researchers have not yet investigated the relative utility of these three personality inventories for predicting positive performance among police officers. The conceptual differences among the inventories and the research findings regarding their utility for predicting performance suggest that each may contribute uniquely to predictive validity in a law enforcement context. As a result, a review of each inventory follows.

THE MMPI

The MMPI is the premier device for screening police officer candidates for psychopathology (Detrick, Chibnall, & Rosso, 2001; Hargrave, 1985; Hargrave, Hiatt, Ogard, & Karr, 1994; Hiatt & Hargrave, 1994; Kornfeld, 1995, 2000; Pallone, 1992). The MMPI yields standardized scores for ten basic clinical scales: 1-hypochondriasis (excessive health concern and somatic focus), 2-depression, 3-conversion hysteria (non-organic somatic complaints), 4-psychopathic deviate (antisocial tendencies, including trouble with the law), 5-masculinity-femininity (gender role confusion), 6-paranoia, 7-psychasthenia (obsessiveness/compulsivity), 8-schizophrenia, 9-hypomania (symptoms of mania), and 0-social introversion (shyness, solitary pursuits, etc.). In addition, there are supplementary scales that measure a variety of constructs, including anxiety, ego strength, family problems, anger, authority problems, and many others. The MMPI-2 also provides three traditional validity scores: the L (Lie) scale, to detect outright faking; the F (Infrequency) scale, to detect random responding or invalid response sets such as exaggerating to gain attention; and the K (Correction) scale, to measure