Concurrent validity of the Candidate and Officer Personnel Survey (COPS)

Robert McGrath‡ and Matthew Guller†

‡School of Psychology T-WH1-01, Fairleigh Dickinson University, Teaneck, NJ 07666, USA. Tel: 201-692-2445; Fax: 201-692-2304; email: mcgrath@fdu.edu
†(Corresponding author) Institute for Forensic Psychology, Suite 2 - 5 Fir Court, Oakland, New Jersey 07436, USA. Tel: 201-337-4996; Fax: 201-337-8378; email: coptest@aol.com

Robert McGrath *is director of the program in clinical psychology and director of the M.S. Programme in Clinical Psychopharmacology at Fairleigh Dickinson University. He is the author of a number of learned articles and books. He has been appointed to several positions, including acting president of the American Psychological Association Division 55 (American Society for the Advancement of Pharmacotherapy); consulting editor to the Journal of Personality Assessment; as a member of the advisory boards for the Research Infrastructure in Minority Institutions Program of Mercy College in Dobbs Ferry, NY; and to the Committee on Science and Practice of the American Psychological Association Division 12 (Society of Clinical Psychology).*

Matthew Guller *is Associate Director and Corporate Counsel of the Institute for Forensic Psychology in New Jersey, USA. He is a licensed attorney, admitted to the bars in New Jersey and Pennsylvania and a licensed psychologist in New Jersey and in New York. In addition to providing psychological services, he oversees legal matters for IFP. Dr. Guller is an associate member of the International Association of Chiefs of Police and is a member of its Legal Services Section. He is also an associate member of the Bergen County, Passaic County and New Jersey State Chiefs of Police Associations. He is an associate member of the American Correctional Association and a contributor to its professional journal.*

ABSTRACT

The Candidate and Officer Personnel Survey (COPS) (Guller & Guller, 2003) was developed as a predictor of job performance in public safety candidates. Three studies were conducted to collect normative data for the COPS, to evaluate the reliability of the scales, and to investigate the concurrent validity of the instrument. Reliability coefficients were generally fair to excellent. Concurrent validity was quite good for the discrimination of promotional versus problematic officers. The results suggest the COPS's potential as a predictor of public safety officers' performance.

INTRODUCTION

A recent survey found that over 90 per cent of police agencies now require psychological testing as part of their selection process (Cochrane, Tett, & Vandecreek, 2003). Though research on the degree to which other categories of public safety candidates such as firefighters undergo psychological evaluation is sparser, there is a substantial body of literature on the topic,

indicating such evaluations are common (eg Barrett, Polomsky, & McDaniel, 1999).

Given the widespread use of psychological measures to predict job performance in public safety candidates, and the significant impact that can result from hiring individuals who are poorly prepared for the work, the ability of psychological instruments to detect potentially problematic candidates is an important research issue. Recently, Varela, Boccaccini, Scogin, Stump, and Caputo (2004) presented the results of a review of the literature on the use of pre-employment personality testing to predict job performance among law enforcement officers. Based on 78 studies that generated almost 4,000 criterion-related validity coefficients, they found the mean correlation was 0.13. After correcting for three attenuation factors — unreliability of predictor variables, range restriction, and dichotomisation of outcomes — this mean increased to 0.22. It is unclear whether they corrected for all dichotomous outcomes or only those that were artificially dichotomised.

Only one inventory in common use has been developed specifically for use with public safety candidates, that being the Inwald Personality Inventory (IPI) (Inwald, Knatz, & Shusman, 1982). Varela et al. (2004) found that across 11 studies that evaluated the effectiveness of the IPI, the mean correlation with criteria was 0.100, which increased to 0.196 after correcting for attenuating factors. These means were lower than those for two instruments not developed specifically for public safety candidates, the California Psychological Inventory (Gough & Bradley, 1996) and the Minnesota Multiphasic Personality Inventory (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989).

The Candidate and Officer Personnel Survey (COPS) (Guller & Guller, 2003) is a 240-item, self-report measure that was developed as an evaluation tool both for individuals seeking employment or promotion as public safety officers, and for officers who have been identified by their jurisdiction for further evaluation because of problematic behaviour. The items were developed to be generally appropriate to police officers as well as firefighters, corrections officers and security personnel. Items are completed on a true-false rather than a polytomous scale in acknowledgement of the diversity in educational backgrounds demonstrated among candidates for these positions.

The COPS generates 18 scales (see Table 1 for a list of scales and their abbreviations) that reflect response bias and positive and negative prediction of work performance. Items were developed initially on intuitive grounds because they reflected some aspect of the construct of interest. Items were modified or eliminated based on a series of small-scale studies of their predictive value.

Scales were not developed with the goal of being homogeneous. In particular, Inconsistency, Overall Prediction, and the LIE scales were developed as formative indices involving items of disparate content. A number of the constructs underlying the scales were also specifically defined as multidimensional constructs, such as Negative Work Attitudes and Bias. Finally, items were chosen primarily for predictive value rather than internal consistency.

The COPS has been used as part of a battery of instruments for the prediction of job performance by public safety candidates for more than 20 years. During this time, the battery as a whole has been administered to more than 40,000 individuals in the United States, Trinidad and Australia who were either seeking employment or promotion as a public safety candidate, or were referred by their agency for testing due to disciplinary or mental health problems. A series of studies has found the battery to be a valid predictor of subsequent job performance in police and corrections

Table 1: Summary of COPS scales and indices

Scale/Index	Scale	Description
Validity Indices		
Lie (LIE)	0–18	Attempt to make a good impression.
Inconsistency (INC)	0–18	Item pairs of essentially equivalent meaning
Positive Indicators		
Success (SUC)	0–132	Diverse items predictive of success
Social Adjustment (ADJ)	0–54	Social adjustment and conformity to social norms
Motivation (MOT)	0–34	Leadership and ambition
Seriousness/Self-Discipline (DIS)	0–42	Self-discipline and self-direction
Overall Prediction (PRE)	31–110	Weighted combination of best success predictors
Negative Indicators		
Alcohol Abuse (ALC)	0–8	Use of and attitudes towards alcohol
Paranoid Orientation (PAR)	0–17	Suspiciousness about others and their motivations
Gender Bias (GEN)	0–11	Attitudes towards women
Personality Problems (PRP)	0–23	Various mental health issues
Depression (DEP)	0–14	Depression
Bias (BIS)	0–10	Racial bias and belief in stereotypes
Authoritarianism (AUT)	0–7	Rigidity and judgmental tendencies
Impulsivity (IMP)	0–15	Impulsive behaviors
Negative Work Attitudes (NEG)	0–23	Negative attitudes about work and work problems
Distrust (DTR)	0–13	Tendency to see others as dishonest
Aggression (AGG)	0–16	Evidence of aggressive tendencies

officers (Fischler, 2004; Guller, 2004; Heyer, 1998; Lough & Ryan, 2005, 2006; Lough, Wald, Byrne, & Walker, 2007). The following presents the results from the first set of studies to be completed evaluating the COPS as a stand-alone instrument.

Study 1

Method

The first study focused on normative data for the instrument. The sample consisted of all 2,509 public safety candidates administered the COPS during the period from January 1994 to December 1996. The sample was largely white males seeking employment as police officers (see Table 2). On average the participants had completed more than one year of college at the time of their testing.

Results and discussion

Table 3 provides basic descriptive statistics for the normative sample. The correlations between the response bias indicators LIE and Inconsistency and other scales were small to moderate, the largest being the negative correlations between LIE and Personality Problems and Depression. The five positive indicators — Success, Social Adjustment, Motivation, Seriousness and Overall Prediction — were all highly correlated with each other, suggesting a strong factor among them. In contrast, all correlations involving at least one negative indicator were substantially smaller, suggesting the potential for greater discriminant validity in the constructs represented. The means and standard deviations at the bottom of Table 3 can be used to generate linear *T* scores.

Study 2

Method

The second study evaluated the reliability of the COPS scales. The sample included 357

Table 2: Demographic statistics for Study 1 (normative statistics)

	N	%	M	SD
Gender				
Male	2264	90.2		
Female	245	9.8		
Cultural Status				
White	1117	74.1		
Black	223	14.8		
Hispanic	146	9.7		
Asian	15	1.0		
Native	7	0.5		
Job Type				
Police	1785	71.1		
Firefighter	278	11.1		
Corrections	153	6.1		
Other[a]	293	11.7		
Age	2508		28.01	6.09
Education	2235		13.68	1.74

Note:

[a] Includes emergency workers and other jobs that involve carrying a gun.

law enforcement personnel and candidates who completed the COPS as part of an evaluation (see Table 4). Again, the majority of candidates were white and male. The sample included 162 job applicants and 195 current public safety officers who were evaluated either as part of the promotional process (21.0 per cent) or because of questions about fitness for duty (33.6 per cent).

Results and discussion

Table 5 provides internal reliability estimates (Cronbach's alpha) for each COPS scale except the three formative measures (Inconsistency, Overall Prediction and LIE). All but two scales (Impulsivity and Negative Work Attitudes) achieved a level of reliability generally considered fair, and six were good to excellent (Fleiss, 1981; Landis & Koch, 1977). Though the scales were not developed with internal consistency in mind, these preliminary analyses suggest that most scales demonstrate evidence of

homogeneity at a level sufficient to indicate that they share an acceptable level of common variability.

Study 3

The third study evaluated the concurrent validity of the COPS using a categorical criterion. COPS profiles were compared for three categories of public safety officers: officers nominated for promotion, officers examined for fitness for duty due to disciplinary issues, and officers examined for fitness for duty due to mental health concerns. It was hypothesised that promotional candidates would demonstrate more positive features than the other two groups, while the mental health group would demonstrate higher scores particularly on scales that are reflective of emotional difficulties.

Method

Nomination for promotion provides an indication of an officer whose performance is superior to that of the typical officer. A sample of 103 (44.8 per cent of the sample) candidates examined for promotion during the period 1998 to 2005 was included in the current study. The sample of officers evaluated for fitness of duty based on disciplinary charges consisted of 74 (32.2 per cent) cases examined during the period 1998 to 2005. Such evaluation could occur because of a single serious infraction, or because of a series of less serious infractions such as abuse of sick time.

The third sample consisted of 53 officers (23.0 per cent) examined for fitness for duty during the period 1998 to 2005 as a result of concerns raised by the agency or the officer regarding possible mental health issues. These included complaints or concerns about significant feelings of stress, anxiety, depression, post-traumatic stress, and substance abuse. In many jurisdictions, a claim of domestic violence involving an officer automatically results in referral for

Table 3: Descriptive statistics for Study 1

	LIE	INC	SUC	ADJ	MOT	DIS	PRE	ALC	PAR	GEN	PRP	DEP	BIS	AUT	IMP	NEG	DTR	AGG
LIE	1.00																	
INC	−0.07*	1.00																
SUC	0.26*	−0.19*	1.00															
ADJ	0.11*	−0.17*	0.81*	1.00														
MOT	0.02	−0.11*	0.60*	0.58*	1.00													
DIS	0.11*	−0.08*	0.62*	0.63*	0.61*	1.00												
PRE	0.01	−0.25*	0.85*	0.84*	0.75*	0.70*	1.00											
ALC	−0.02	0.07*	−0.29*	−0.25*	−0.12*	−0.20*	−0.28*	1.00										
PAR	−0.16*	0.20*	−0.51*	−0.42*	−0.36*	−0.18*	−0.45*	0.13*	1.00									
GEN	−0.20*	0.17*	−0.26*	−0.18*	−0.12	−0.07	−0.27*	0.05	0.40*	1.00								
PRP	−0.31*	0.14*	−0.66*	−0.47*	−0.25*	−0.27*	−0.45*	0.22*	0.44*	0.34*	1.00							
DEP	−0.37*	0.16*	−0.59*	−0.40*	−0.29*	−0.24*	−0.41*	0.14*	0.46*	0.31*	0.74*	1.00						
BIS	−0.02	0.16*	−0.41*	−0.33*	−0.29*	−0.17*	−0.38*	0.16*	0.54*	0.48*	0.36*	0.34*	1.00					
AUT	−0.11*	0.13*	−0.36*	−0.26*	−0.20*	−0.10*	−0.32*	0.13*	0.56*	0.36*	0.32*	0.34*	0.51*	1.00				
IMP	−0.17*	0.10*	−0.43*	−0.44*	−0.22*	−0.32*	−0.44*	0.40*	0.22*	0.07	0.27*	0.25*	0.22*	0.22*	1.00			
NEG	−0.12*	0.11*	−0.59*	−0.49*	−0.74*	−0.50*	−0.65*	0.13*	0.35*	0.15*	0.32*	0.30*	0.25*	0.23*	0.22*	1.00		
DTR	−0.28*	0.16*	−0.53*	−0.37*	−0.29*	−0.24*	−0.42*	0.15*	0.74*	0.37*	0.43*	0.47*	0.50*	0.60*	0.31*	0.34*	1.00	
AGG	−0.27*	0.14*	−0.38*	−0.39*	−0.23*	−0.21*	−0.41*	0.40*	0.31*	0.21*	0.29*	0.23*	0.19*	0.29*	0.47*	0.26*	0.31*	1.00
M	6.76	1.85	109.82	48.65	25.77	25.97	55.47	0.63	2.15	0.44	0.72	1.08	0.75	2.62	1.84	4.72	3.52	2.62
SD	3.31	1.07	7.23	3.74	3.11	3.73	23.51	0.83	1.89	0.88	1.92	1.42	1.19	1.75	1.31	1.96	2.32	1.39

Notes:
*p < 0.05
N = 2,509 except for GEN and AGG (N = 271) and PRE (N = 2311), which were added later in the development process.
Abbreviations may be found in Table 1.

Table 4: Demographic statistics for Study 2 (internal consistency)

	N	%	M	SD
Gender				
Male	338	94.7		
Female	19	5.3		
Cultural Status				
White	224	84.5		
Black	13	4.9		
Hispanic	24	9.1		
Asian	3	1.1		
Native	1	0.4		
Reason for Evaluation				
Candidate	162	45.4		
Promotion	75	21.0		
Fitness	120	33.6		
Age	357		32.44	8.44
Education	357		14.09	1.90
Tenure	195		12.59	7.54

Note:

Candidate = candidate for employment; Promotion = candidate for promotion; Fitness = fitness for duty in question; Tenure = length of employment (years).

Table 5: Internal consistency estimates

Scale/Index	Cronbach's Alpha
Success (SUC)	0.830
Social Adjustment (ADJ)	0.657
Motivation (MOT)	0.522
Seriousness/Self-Discipline (DIS)	0.560
Alcohol Abuse (ALC)	0.529
Paranoid Orientation (PAR)	0.723
Gender Bias (GEN)	0.646
Personality Problems (PRP)	0.831
Depression (DEP)	0.803
Bias (BIS)	0.729
Authoritarianism (AUT)	0.550
Impulsivity (IMP)	0.307
Negative Work Attitudes (NEG)	0.396
Distrust (DTR)	0.681
Aggression (AGG)	0.544

mental health evaluation. Such cases were excluded from this group due to possible concerns about the veracity of the claim, and because claims of domestic abuse generally did not involve any job-related problems.

Cases were selected at random from existing archives for the three groups. They were excluded from the disciplinary and mental health groups when the clinical team conducting the evaluation (consisting of two clinical psychologists) concluded that the officer was malingering. Given the clinical setting in which these evaluations took place, no reliability data were available concerning these judgments.

Table 6 summarises demographic information for the sample as a whole and for each of the three groups. A significant difference was found in gender, χ^2 (2, $N = 230$) = 11.9, $p = 0.003$. Though all three groups were predominantly male, the proportion of females was highest among those undergoing mental health evaluations, and lowest among the promotional candidates. Promotional candidates also reported an average of one more year of education than participants undergoing a disciplinary or mental health evaluation, $F(2, 227) = 6.2$, $p = 0.002$. They were on the job for significantly longer than the disciplinary cases, $F(2, 226) = 6.6$, $p = 0.002$, with those undergoing mental health evaluations falling in between and not significantly different from either.

In addition to the COPS, data were available from the Shipley Institute of Living Scale (SILS) (Zachary, 2000), which is a widely used, brief measure of intellectual functioning. Given a substantial literature demonstrating the importance of cognitive ability as a predictor of job performance (Ones, Viswesvaran, & Dilchert, 2005), the concurrent validity of the SILS offered a standard by which to compare the effectiveness of the COPS. The results of the Varela

Table 6: Demographic statistics for Study 3 (concurrent validity)

	N	%	Promotional		Disciplinary		Mental Health	
			N	%	N	%	N	%
Gender★								
Male	215	93.5	102	99.0	68	91.9	45	84.9
Female	15	6.5	1	1.0	6	8.1	8	15.1
Cultural Status								
White	119	51.7	53	94.6	33	86.8	33	97.1
Black	3	1.3	0	0.0	3	7.9	0	0.0
Hispanic	5	2.2	2	3.6	2	5.3	1	2.9
Asian	1	0.4	1	1.8	0	0.0	0	0.0
	M	SD	M	SD	M	SD	M	SD
Age	38.12	7.01	38.37	6.79	37.88	6.70	37.96	7.92
Education★	13.93	1.92	14.41[a]	1.97	13.50[b]	1.69	13.58[b]	1.93
Tenure★	12.03	7.04	13.81[a]	6.87	10.13[b]	6.43	11.24	7.47

Notes:

★Omnibus test comparing groups significant at $p < 0.05$

If the omnibus test for a quantitative variable is significant, means with different superscripts are significantly different according to Tukey tests ($p < 0.05$). Tenure = length of employment.

et al. (2004) meta-analysis provided a second basis for comparison.

Results

Analysis of variance results is provided in Table 7 for each COPS scale as well as for the SILS. All but two omnibus tests for the COPS scales (Inconsistency and Authoritarianism) were significant, though two others demonstrated no differences on post-hoc Tukey tests (LIE and Gender Bias). A consistent pattern emerged for the positive indicators (Success, Social Adjustment, Motivation, Seriousness, and Overall Prediction) of significant differences between the promotional group and the other two groups but no difference between the disciplinary and mental health groups. *d* values for pairwise comparisons involving the promotional group were substantial, suggesting group means differed by one standard deviation or more. These effect sizes were much larger than those observed for the

SILS. Point-biserial correlation values are also reported to allow direct comparison with Varela et al.'s (2004) findings. Specifically, Varela et al. found an average uncorrected correlation for concurrent validity analyses of 0.20; the mean correlation comparing promotional candidates to another group in the present study was 0.49, which is approximately a six-fold increase in the proportion of variance accounted for.

Varela et al. (2004) undoubtedly combined studies where group base rates were allowed to vary freely with studies where group base rates were artificially equalised. In contrast, the base rates for the three groups in the present study reflected their relative frequency within the population that has completed the COPS, a procedure that could potentially result in a lower value for the correlations. However, using a formula to control for the effect of unequal base rates on the correlation coefficient (see McGrath & Meyer, 2006, Table 3),

Table 7: Test outcomes by group

	Promotional (P)		Disciplinary (D)		Mental Health (M)		d			r		
	M	SD	M	SD	M	SD	P-D	P-M	D-M	P-D	P-M	D-M
SILS★	107.50[a]	5.98	102.38[b]	9.96	105.34	5.74	0.65	0.37	-0.35	0.30	0.17	-0.17
LIE★	4.33	2.79	5.34	3.23	4.09	2.84	-0.34	0.08	0.40	-0.16	0.04	0.20
INC	1.50	0.84	1.84	1.23	1.57	1.07	-0.34	-0.08	0.23	-0.16	-0.04	0.11
SUC★	112.28[a]	5.79	103.15[b]	8.81	102.36[b]	9.43	1.27	1.37	0.09	0.53†	0.54†	0.04
ADJ★	50.29[a]	3.10	46.04[b]	4.07	46.79[b]	4.83	1.20	0.93	-0.17	0.51†	0.40†	-0.08
MOT★	30.25[a]	2.48	26.97[b]	3.42	27.02[b]	3.58	1.13	1.11	-0.01	0.49†	0.47†	-0.01
DIS★	26.95[a]	3.39	23.38[b]	3.86	24.26[b]	4.25	0.99	0.73	-0.22	0.44†	0.33†	-0.11
PRE★	76.21[a]	21.75	38.15[b]	30.65	35.77[b]	33.99	1.47	1.52	0.07	0.59†	0.59†	0.04
ALC★	0.48[a]	0.65	1.01[b]	1.24	1.64[c]	1.97	-0.57	-0.92	-0.40	-0.27†	-0.40†	-0.19†
PAR★	2.09[a]	1.66	2.43	2.40	3.11[b]	2.55	-0.17	-0.51	-0.28	-0.08	-0.24†	-0.14
GEN★	0.88	1.17	0.85	1.29	1.38	1.48	0.03	-0.38	-0.38	0.01	-0.18†	-0.19†
PRP★	1.80[a]	2.02	3.34[b]	3.94	5.11[c]	3.73	-0.52	-1.22	-0.46	-0.25†	-0.50†	-0.22†
DEP★	1.17[a]	1.37	2.11[b]	2.48	3.49[c]	2.89	-0.49	-1.15	-0.52	-0.24†	-0.48†	-0.25†
BIS★	0.41[a]	0.91	0.49[a]	0.94	0.98[b]	1.74	-0.09	-0.46	-0.37	-0.04	-0.21†	-0.18†
AUT	2.06	1.39	2.05	1.38	2.58	1.91	0.00	-0.33	-0.33	0.00	-0.16	-0.16†
IMP★	1.01[a]	1.48	1.95[b]	1.71	2.25[b]	1.59	-0.59	-0.81	-0.18	-0.28†	-0.36†	-0.09
NEG★	3.51[a]	1.69	6.04[b]	2.49	5.98[b]	2.65	-1.23	-1.20	0.02	-0.52†	-0.49†	0.01
DTR★	3.31	2.02	3.00[a]	2.09	4.15[b]	2.47	0.15	-0.38	-0.51	0.07	-0.18†	-0.24†
AGG★	2.45[a]	1.41	3.19[b]	2.08	3.23[b]	1.83	-0.43	-0.50	-0.02	-0.21†	-0.23†	-0.01

Notes:

★Omnibus test significant at $p < 0.05$. †Significantly predicts over SILS at $p < 0.05$ based on simultaneous logistic regression.

If the omnibus test for a quantitative variable is significant, means with different superscripts are significantly different according to Tukey tests ($p < 0.05$).

SILS = Shipley Institute for Living Scale; all other abbreviations may be found in Table 1.

the mean of these 10 correlations only increased to 0.50.

In almost every case there was a significant difference in the expected direction between the promotional group and at least one of the other groups. Finally, for each pairs of groups a binary variable was created and regressed onto the SILS and each COPS scale individually via simultaneous logistic regression. Even after accounting for variability due to differences in intellectual performance, all five positive predictors distinguished significantly between the promotional and the other two groups ($p < 0.05$). The consistency of differences was particularly striking when the promotional candidates were compared with the mental health group: every substantive COPS scale predicted significantly over intellectual performance except for Authoritarianism.

As predicted, the mental health group was most distinguishable from the other groups on the basis of scales associated with emotional difficulties (Alcohol Abuse, Paranoid Orientation, Personality Problems, and Depression), though these differences were small to moderate in size. In each case, the effect was larger when the mental health group was compared with the promotional group than with the disciplinary group. The results consistently and strongly supported the validity of the COPS as a concurrent predictor of categorical status.

DISCUSSION

This preliminary investigation suggests that, while the reliability of some scales could be improved, the COPS was an effective predictor of whether a candidate was being seen for promotion or for job difficulties. In some cases the relationships were substantially larger than those reported for a measure of cognitive performance, and the COPS demonstrated incremental validity for distinguishing between groups. The results indicate a promising method of evaluating public safety officers. It will be important for future research to focus on establishing the predictive validity of the COPS for prospective candidates.

Despite widespread use of psychological testing in the assessment of public safety officers, the existing evidence indicates there is substantial room for improvement. To some extent, low validity coefficients reported by Varela et al. (2004) may be due to range restriction, both because many questionable candidates are screened out before they are referred for psychological assessment, and because the psychological screening itself rules out additional candidates. However, even after correcting for range restriction the results for measures in common use today fall at the low end of the acceptable range. This finding may partly reflect a failure to develop indicators that are specific to the issues relevant to effective performance as a public safety candidate today, such as issues of racial or gender bias. The COPS potentially offers an approach to addressing these gaps in the available tools.

NOTE

REFERENCES

Barrett, G. V., Polomsky, M. D., & McDaniel, M. A. (1999). Selection tests for firefighters: A comprehensive review and meta-analysis. *Journal of Business and Psychology, 13*, 507–513.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI-2: Manual for administering and scoring.*

Minneapolis, MN: University of Minnesota Press.

Cochrane, R. E., Tett, R. P., & Vandecreek, L. (2003). Psychological testing and the selection of police officers: A national survey. *Criminal Justice and Behavior, 30,* 511–537.

Fischler, G. L. (2004). *Identifying psychological predictors of police officer integrity problems.* Washington, DC: Community Oriented Policing Services Office of the United States Department of Justice.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions.* New York: Wiley.

Gough, H. G., & Bradley, P. (1996). *CPI manual* (3rd ed.). Mountain View, CA: CPP.

Guller, M. (2004). *Predicting performance of law enforcement personnel using the Candidate and Officer Personnel Survey and other psychological measures.* West Orange, NJ: Seton Hall University.

Guller, I. B., & Guller, M. (2003). *C.O.P.S.: Candidate and Officer Personnel Survey technical manual.* Oakland, NJ: Institute for Forensic Psychology.

Heyer, T. (1998). *A follow-up study of the prediction of police officer performance on psychological evaluation variables.* Minneapolis, MN: Minnesota School of Professional Psychology.

Inwald, R. E., Knatz, H., & Shusman, E. (1982). *Inwald Personality Inventory manual.* Kew Gardens, NY: Hilson Research.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174.

Lough, J., & Ryan, M. (2005). Psychological profiling of Australian police officers: An examination of post-selection performance. *International Journal of Police Science & Management, 8,* 15–23.

Lough, J., & Ryan, M. (2006). Psychological profiling of Australian police officers: A longitudinal examination of post-selection performance. *International Journal of Police Science & Management, 8,* 143–152.

Lough, J., Wald, E., Byrne, K., & Walker, G. (2007). The impact of psychological profiling of Australian correctional officers. *Corrections Compendium, 32*(4), 1–6.

McGrath, R. E., & Meyer, G. J. (2006). When effect sizes disagree: The case of *r* and *d*. *Psychological Methods, 11,* 386–401.

Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Cognitive ability in selection decisions. In: O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 431–468). Thousand Oaks, CA: Sage.

Varela, J. G., Boccaccini, M. T., Scogin, F., Stump, J., & Caputo, A. (2004). Personality testing in law enforcement employment settings: A meta-analytic review. *Criminal Justice and Behavior, 31,* 649–675.

Zachary, R. A. (2000). *Shipley Institute of Living Scale: Revised manual.* Los Angeles: Western Psychological Services.