Validity Study

Relationship of Cognitive, Biographical, and Personality Measures with the Training and Job Performance of Detention Enforcement Officers in a Federal Government Agency

Thomas J. Lyons, J. Anthony Bayless, & Randolph K. Park
U.S. Immigration and Naturalization Service

Scores on a predictor battery of cognitive, biographical, and personality measures completed by a nationwide sample of 381 Detention Enforcement Officers were correlated with three criterion measures: the officers' training scores, work simulation scores, and supervisory ratings of job performance. All predictor measures were significantly correlated with all criterion measures (uncorrected r's ranged from .13 to .57). However, the personality measure did not provide incremental validity in predicting the composite criterion when combined with the cognitive and biographical measures in a stepwise multiple regression analysis. A unit-weighted composite of cognitive and biographical measures was proposed for the test battery, and it correlated .59 with the composite criterion.

Sample

N	381 Federal Detention Enforcement Officer (DEO) employees
Gender	92% were men, 8% were women
Race	42% were White, 11% were African American, 42% were Hispanic, 4% were Asian, and 1% were American Indians
Age	Mean = 35.6, Standard deviation = 6.1

Location

The data for this study were collected at Federal detention facilities and field offices at 14 sites throughout the continental United States in 2001.

Predictor Information

The predictors in the study consisted of three cognitive tests (Name and Number Comparison, Logical Reasoning, and Following Policies and Procedures), two biographical data measures, and an Assessment Inventory. Name and Number Comparison is a speeded test consisting of 50 items. Names or numbers from a list must be compared with five choices, one of which is the correct match. The Logical Reasoning test has 30 questions relating to short paragraphs containing several logical schemas. It is a measure of deductive reasoning. The Following Policies and Procedures

test has 20 questions and measures the ability to read and apply organizational policies and procedures. One biographical measure was developed for selecting Federal employees and contains items related to school, work, and personal experiences. It has an empirical key based on incumbents in professional and administrative jobs. The second biographical measure contains questions about school and work experience. It has an empirical key based on state and local-government law enforcement occupations. The Assessment Inventory contains biographical data, temperament items, and situational judgment questions. It has scales for conscientiousness, cooperativeness, emotional maturity, initiative, integrity/honesty, and judgment, one scale for situational judgment, and an overall score. It has a rationally developed scoring key based on construct validity evidence. Table 1 contains the reliabilities and intercorrelations of the six predictors.

Table 1
Predictor Reliabilities and Intercorrelations

Predictor	(1)	(2)	(3)	(4)	(5)	(6)
Cognitive Ability						
Name and number comparison (1)	(.82)	.55	.54	.24	.19	.14
Logical reasoning (2)		(.82)	.69	.24	.21	.19
Following policies and procedures (3)			(.71)	.21	.17	.19
Biodata						
Entry-level federal key (4)				(.88)	.74	.52
Law enforcement key (5)					(.75)	.58
Personality						
Assessment Inventory (6)						(.94)

Note: Reliability estimates are in parentheses. All coefficients are significant at the .01 level.

Criterion Information

The criteria were training scores, work simulation scores, supervisory ratings of job performance, and a unit-weighted composite. The training score was a composite of the final grades from courses in immigration law and police procedures that DEO trainees completed at the training academy. The work simulation score was the number of correct answers to questions based on written work-related scenarios with four response alternatives that described possible courses of action. The supervisory rating score was a composite of duty- and competency-based ratings on scales that were anchored with behavioral statements of job performance. Table 2 contains the reliabilities and intercorrelations of the four criteria.

Validity Information

Table 3 presents the correlations of the predictor measures with criterion measures for this concurrent validity study. A stepwise multiple regression analysis was run on the predictor measures against the composite criterion. Significant positive beta weights were obtained for the Logical Reasoning test, Following Policies and Procedures

test, and law enforcement biodata measure. Rounding the beta weights to whole numbers resulted in a unit-weighted composite that correlated .65 (uncorrected) with the composite criterion. For this predictor composite the effect size was .80 for African Americans and .88 for Hispanics in a simulated applicant population. The effect size is the mean score difference between a minority group and Whites described in standard deviation units. A positive effect size indicates that Whites are performing higher than a minority group on the predictor. Although these effect sizes were below the levels commonly observed for cognitive ability tests, a lower effect size was considered desirable for operational use.

Table 2
Criterion Reliabilities and Intercorrelations

Criterion	Training Performance	Work Simulation	Supervisor Ratings	Composite Score
Training performance	(.80)	.30**	.19**	.73**
Work simulation		(.80)	.09	.68**
Supervisory ratings of job performance			(.60)	.63**
Composite score				(.81)

Note: Reliability estimates are in parentheses. **$p<.01$

Table 3
Validity Coefficients

Predictor Measures	Criterion Measures			
	Training Performance	Work Simulation	Supervisory Rating	Composite Score
Name and Number Comparison	.38 (.34)	.27 (.24)	.25 (.19)	.42 (.38)
Logical Reasoning	.58 (.52)	.48 (.43)	.28 (.21)	.64 (.57)
Following Policies and Procedures	.56 (.50)	.48 (.43)	.22 (.17)	.60 (.54)
Biodata – entry-level Federal key	.24 (.21)	.20 (.18)	.20 (.15)	.29 (.27)
Biodata – law enforcement key	.25 (.23)	.23 (.21)	.37 (.29)	.39 (.35)
Assessment Inventory – total score	.14 (.13)	.19 (.17)	.18 (.14)	.24 (.21)
Highest Unit Weighted Composite	.61 (.55)	.53 (.47)	.38 (.29)	.72 (.65)
Proposed Unit Weighted Composite	.52 (.47)	.43 (.38)	.30 (.23)	.59 (.53)

Note: Coefficients are corrected for unreliability in the criteria. Coefficients in parentheses are uncorrected. All uncorrected coefficients are statistically significant at the .05 level or higher.

Other combinations of predictor measures were explored to identify one with a lower effect size. A unit-weighted composite of the Logical Reasoning test and the Federal biodata inventory correlated .53 (uncorrected) with the composite criterion. This predictor composite had a much lower effect size for African Americans (.57) and Hispanics (.58) than the previous composite, with a small reduction in validity. Thus it was proposed for operational use.

Although the scores on composite predictor and criterion measures were significantly higher for Whites than for African Americans or for Hispanics, results of a fairness analysis based on the equivalence of subgroup regression equations indicated that the composite measure was fair for Hispanics. The sample size for Africans Americans was too small (n<50) to conduct fairness analysis for this subgroup. There were no significant differences between males and females in their performance on predictor or criterion measures for the proposed model.

In the present study, the Assessment Inventory, which has six personality scales and a Situational Judgment scale, did not provide incremental validity beyond cognitive and biodata measures. Table 4 presents the correlation of each scale with the criterion measures. Four personality scales and the Situational Judgment scale were significantly correlated with the composite criterion. When biodata measures were excluded from a stepwise multiple regression analysis of predictor measures with the composite criterion, two scales (Initiative and Situational Judgment) provided significant incremental validity. These findings suggest that personality and situational judgment measures might be a useful addition to a cognitive test for predicting law enforcement performance in the absence of biodata measures.

Table 4
Validity Coefficients (uncorrected)

Assessment Inventory Scales	Criterion Measures			
	Training Performance	Work Simulation	Supervisory Rating	Composite Score
Conscientiousness	.03	.18*	.10	.16*
Cooperativeness	.11*	.11*	.11*	.16**
Emotional Maturity	.12*	. 04	.16**	.16**
Initiative	.20**	.18**	.18**	.27**
Honesty/Integrity	.01	-.04	.02	-.01
Judgment	.05	.04	.07	.08
Situational Judgment	.14**	.22**	.24**	.29**

*$p<.05$, **$p<.01$

Questions about this validity study should be addressed to:
Thomas J. Lyons, Ph.D.
U.S. Immigration and Naturalization Service
800 K Street, NW, Room 5000
Washington DC 20536
(202) 305-4884
thomas.j.lyons@justice.usdoj.gov