

AD-A267 502 DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

This is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and reviewing the collection of information, sending comments regarding this burden estimate or any other aspect of this burdening this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Avenue, Suite 1204, Alexandria, VA 22304-6145, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE Aug 1993	3. REPORT TYPE AND DATES COVERED THESIS/DISSERTATION	
4. TITLE AND SUBTITLE Development and Validation of the University of Texas System Entry-Level Police Examination		5. FUNDING NUMBERS XXXXXX	
6. AUTHOR(S) Anne C. Parker			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AFIT Student Attending: The University of Texas At Austin		8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/CI/CIA-93-042	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) DEPARTMENT OF THE AIR FORCE AFIT/CI 2950 P STREET WRIGHT-PATTERSON AFB OH 45433-7765		10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release IAW 190-1 Distribution Unlimited MICHAEL M. BRICKER, SMSgt, USAF Chief Administration		12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words)			
14. SUBJECT TERMS		15. NUMBER OF PAGES 109	
		16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

DEFENSE TECHNICAL INFORMATION CENTER



9318122

DEVELOPMENT AND VALIDATION OF THE UNIVERSITY OF TEXAS
SYSTEM ENTRY-LEVEL POLICE EXAMINATION

APPROVED:

Joseph M. Horn
Robert L. Helms

Copyright
by
Anne Caroline Parker
1993

Abstract

DEVELOPMENT AND VALIDATION OF THE UNIVERSITY OF TEXAS SYSTEM ENTRY-LEVEL POLICE EXAMINATION

by

Anne Caroline Parker, First Lieutenant, United States Air Force

1993

109 pages

Master of Arts

The University of Texas at Austin

The experimenter developed and validated a new entry-level police examination for The University of Texas System Police (UTSP). The predictor in the study was the selection test, and the criteria in the study were UTSP Academy performance averages. Both cadets at the Academy and incumbents took the entry-level test. Their test scores were correlated with their performance averages, which enabled the experimenter to perform a criterion-related validity study. Both concurrent and predictive validation methods were utilized. There were no significant differences between scores from different race and gender groups. Results indicated the test was an accurate and fair predictor of performance at the UTSP Academy. The entry-level test provided a solid foundation for the selection of competent campus police officers.

Author Notes

I gratefully acknowledge the assistance of the Office of the Director of Police, The University of Texas System, for supporting this research. I especially wish to thank Inspector Lawrence E. Coutorie for coming to my rescue during innumerable computer crises.

I wish to express my gratitude to my advisor Dr. Joseph M. Horn for his guidance and support in preparing this thesis. His knowledge and advice have made the completion of this project possible.

I also wish to thank Dr. Robert L. Helmreich and Kate Morris for comments on several drafts of this project. Most of all I extend my sincere thanks to my future husband, Tom Duke, for his unwavering support and encouragement.

Table of Contents

Author Notes	iv
Introduction	1
Method	28
Results	40
Discussion	49
References	59
Appendices	66
Tables	99
Vita	110

List of Tables

The University of Texas System Police Departments That Participated in the Development Phase of the Entry-Level Examination	99
The Linkage Between the Eight Cognitive Abilities and the Subtest Areas of the Entry-Level Examination	100
The University of Texas System Police Departments That Provided Incumbents for Testing	101
Race and Gender Compositions of the Cadet Sample, Incumbent Sample, and the UTSP Force	102
Means, Variances, and Correlations Among the Dependent Variables and the Independent Variables	103
Test Statistics (Main II Test and Refugio Test)	105
Correlation Matrix (Incumbent Sample)	106
Test Score Differences by Race, Sex, and Group	107
Attrition Summary for UTSP Academy Classes 46 Through 55	108
Race and Gender of the Resignees from UTSP Academy Classes 46 Through 53 (Compared to UTSP Composition Percentages)	109

Development and Validation of The University of Texas System Entry-Level Police Examination

Purpose

Organizations require a foundation of competent personnel in order to achieve their objectives. In order to acquire a strong personnel foundation, organizations must carefully select and train their employees. A good personnel selection procedure identifies individuals who will make constructive contributions to the organization. At the same time, a good selection device screens out those applicants who will not make positive inputs to the organization.

Adequate selection is one of the most troublesome, expensive, and time-consuming tasks facing organizations (Colarelli & Siegel, 1964). The skillful selection of employees is especially important in the law enforcement domain. According to Cascio (1987), the entire police selection process including recruiting, training, salaries, uniforms, and other items may cost as much as \$50,000 or more per officer candidate.

The importance of an effective and efficient police force to any community is self-evident. The reliable administration of criminal justice depends upon police officers' courage, skill, and intelligent application of the law (Blum, 1964). Usually there are multitudes of applicants for each vacancy on

a police force. Therefore, the written test, which is usually the first stage of the selection procedure, is crucial in its role of screening candidates from an applicant pool. A properly developed and validated test aids police departments in the pre-hiring identification of those applicants who are most likely to be capable and effective on the job.

The University of Texas System Police (UTSP) desired an updated selection tool. In 1992 the UTSP coordinated with the University of Texas at Austin (UT-Austin) to develop a new entry-level police officer examination. The updated test would help to bring two opposing forces together. One force, which tried to keep standards low, set criteria for employment which did not eliminate a large proportion of applicants. Behind this goal was the desire of police departments to minimize delays in filling vacancies (Blum, 1964). In conflict with the emphasis on low standards there was an increasing force in the UTSP to set higher standards. With higher standards, fewer applicants can meet the criteria. When fewer applicants pass validated selection tests, there are correspondingly fewer future employees who perform poorly on the job.

The purpose of this study was to investigate the relationship between performance on an entry-level police officer test and performance in the UTSP Academy. Clearly,

this project would not have been possible without the cooperation and participation of all UTSP departments.

Major goals of the project were as follows:

1. To perform a job analysis of the UTSP across all 15 campuses;
2. To develop an entry-level police officer test that assessed the characteristics (amenable to the constraints of written testing procedures) required of police recruits;
3. To validate the test as a consistent predictor of performance at the UTSP Academy;
4. To analyze the test for its compliance with legal and professional guidelines and standards;
5. To ensure the fairness and the job applicability of the test for all groups, with a special focus on minorities and females.

Utility

Utility is an important concept when considering the cost and time effectiveness associated with a selection test. The utility of a selection test includes the accuracy of the decisions made, the costs of using the selection system, and the costs incurred by the errors involved in the decision process (Arvey & Faley, 1988; Zeidner & Johnson, 1991).

Dunnette (1966) categorizes the different types of costs into actual and potential costs. Actual costs include those

incurred when an applicant is hired (Dunnette, 1966). To develop and implement a selection tool requires much time and funding. Funds are also required to provide training and professional development for the officer. For this reason, early identification of those candidates most likely to complete training reduces the number of cadets and results in actual cost benefits to the campus.

Employee selection can also incur potential or indirect costs. Potential costs are those associated with faulty selection decisions (Dunnette, 1966). Smart (1983) estimates that hiring the wrong individual for a position can cost three-and-one-half times the individual's salary. One of the potential costs is the waste of money involved in accepting candidates who do not later meet the police department's performance standards. Hiring substandard candidates not only results in wasted training expenses and poor services rendered, but damage to the reputation of the campus police. Unsatisfactory job performance can be detrimental if not dangerous to the campus community and fellow police officers. Therefore, campuses benefit from early grouping of those candidates who are likely to provide competent service.

Discrimination and Legislation

While avoiding both actual and potential costs, the basic objective of a selection test is to discriminate between unfit

applicants and suitable applicants. Such pre-employment evaluation "screens out" unfit applicants and "selects in" individuals who will most likely adapt to the job stresses and demands such as those found in police work (Fabricatore, 1979). The more selective a test is, the more increases one can expect in the performance level of those applicants selected (Martin & Raju, 1992). For these reasons, according to Guion (1966), all employment practices are, by nature, discriminatory.

It is noteworthy that there is considerable confusion between discrimination in a legal sense and discrimination in a statistical sense (Arvey & Faley, 1988). "Discriminating" in statistical context refers to providing a means by which some applicants may be selected and others may be rejected. A critical parameter affecting the ability of a test to "discriminate" in a statistical sense is variability. All useful selection tests require enough variability among the applicants in order to make discrimination possible. A selection tool that does not demonstrate much variability among applicants provides little or no useful information with which to make an employment decision.

In a legal sense, the topic of "discrimination" in employment selection has negative connotations. The following definition of "unfair discrimination" constructed by

Arvey and Faley (1988) provides the reader with an understanding of the legal interpretation of unfair discrimination:

Unfair discrimination or bias is said to exist when members of a minority group have lower probabilities of being selected for a job when, in fact, if they had been selected, their probabilities of performing successfully on the job would have been equal to those of non-minority group members. (p. 7)

When developed and implemented properly, it is obvious that selection methods have clear benefits to police departments. Paying careful attention to prevent unfair discrimination facilitates the realization of these benefits.

It is difficult for organizations such as the UTSP to cope with legislative trends because fixed budgets usually do not provide ample funding for the development of updated testing instruments. However, the International Personnel Management Association (IPMA, 1988) warns that the cost of litigation and settlement can far exceed the costs of developing new tests. Therefore, organizations must attend to legislative events that restrict employment practices.

There are three such legislative events which have drastically affected employee selection practices in police departments as well as in other public and private

organizations (Feild, Holley, & Buckner, 1990). The formation of the Equal Employment Opportunity Commission (EEOC) has resulted in the creation of strict employment selection rules. The passage of the Uniform Guidelines on Employee Selection Procedures (EEOC, 1978) and the Civil Rights Act of 1991 has also raised concern in employers.

The Civil Rights Act of 1991 did not create many new rights in the workplace. It also did not establish any new requirements regarding changes in employers' personnel practices (Grider & Bailey, 1992). The new law dealt almost entirely with the alteration of the technical rules for trying employment discrimination cases in court (Brown Maroney, 1991).

The courts developed the *disparate impact* theory of discrimination under Title VII of the Act. The theory focused on the comparison between selection rates for majority and minority groups (Sackett & Arvey, 1993). According to the theory, an individual could challenge a putatively neutral selection device on the grounds that it had a disproportionate impact on minorities or that it was not related to performance on the job (Morrison & Foerster, 1992; Moynihan, 1992). Such adverse or disparate impact was said to occur when an employer applied a selection practice equally to all applicants, but the practice fell more harshly on a protected group than on

others (Feild et al., 1990; Moynihan, 1992). In order to refute a challenge, a plaintiff could prove that a particular selection practice had business justification or "business necessity" (Morrison & Foerster, 1992; Sauls, 1992).

Additionally, Section 106 of the Act prohibits the practice of "race-norming" in employment selection tests. "Race-norming" refers to the practice of adjusting scores or setting different cutoff scores on the basis of race, color, sex, religion, or national origin (Grider & Bailey, 1992; Morrison & Foerster, 1992; Sauls, 1992). In other words, the Act prohibits the use of setting different within-group test norms to eliminate adverse impact while retaining most of the productivity benefits associated with selection (Schmidt, 1993).

Besides being wary of a selection test's potential for discriminatory impact, employers must also assess the job-relatedness of the selection test itself (Gatewood & Feild, 1990). The Uniform Guidelines (EEOC, 1978) dictate the means by which an employer can prove a selection device has such job relatedness. These guidelines ensure that an employment selection test demonstrates validity. The employment standards set by the EEOC require employers to validate the process by which they select employees (Jurkanin, 1989). The validation process entails determining whether a specific

standard measures the knowledge, skills, and abilities (KSAs) it is designed to assess.

While requesting a test that would perform adequate selection, the UTSP also wanted a test that would adhere to current legal specifications. The specifications in the Civil Rights Act of 1991 and the Uniform Guidelines (EEOC, 1978) were taken into account during the development and validation of the UTSP entry-level police officer examination.

Job Analysis

Developing a selection measure requires the execution of a job analysis (Jurkanin, 1989). According to Sackett and Arvey (1993), there are three main purposes for which job analyses are performed in the context of selection. They are development of criteria, development of predictors, and determination of job similarity.

Criteria Development.

The UTSP has already developed the criteria for its police officers enrolled in the UTSP Academy. The UTSP's criteria development enabled the test developers to determine which components of job performance were important. It also facilitated the development of measures of the important components.

The researchers were able to define the job of a campus law enforcement officer in terms of tasks and duties. Lists of

tasks and duties were available from a recent analysis conducted by IPMA (1992). IPMA's (1992) list labeled the following categories of duties: accident scene duties, apprehension and arrest, citizen interaction, court preparation and testimony, crime scene duties, handling disturbances, investigation and interviewing, patrol duties, reports and paperwork, self-development, special duty assignments, and station duties. The UTSP Headquarters agreed that the subjects taught in the UTSP Academy and the requirements on the job were comparable to those listed by IPMA.

Predictor Development.

The second purpose of job analysis data is to develop predictors (Sackett & Arvey, 1993). Predictor construction can include the development of candidate interviews, personality tests, assessment center exercises, and job knowledge tests. All of the predictors are devised to sample factors within a job domain. Job analysis data help specify that domain. In order to develop test items, all UTSP agencies cooperated. All campuses provided valuable input for the development of the test. The list of campus police departments that participated in the test development phase appears in Table 1.

Insert Table 1 about here

Because job-related tasks are not easily assessed by pencil-and-paper tests, task-oriented job analyses have often created obstacles for test-writing. Since the purpose of this study was to develop a written examination, the researchers have expanded the job analysis. According to Thorndike (1962), job and trait approaches are merely extremes on a continuum, not unrelated approaches. Arvey and Faley (1988) add that it is common to use a mixture of job analysis procedures.

The trait approach which Thorndike (1962) addresses refers to the ability-oriented job analysis procedure. Such a procedure involves the collection of information about the abilities needed to perform a job. This procedure emphasizes the KSAs that are required for effective performance (Arvey & Faley, 1988). An ability-oriented approach is desirable in the case of the UTSP since the selection test is to be used for the selection of entry-level police officers. The test cannot require prior job exposure; therefore, it must be based primarily on abilities and aptitudes rather than on performance of actual job behaviors.

The ability-oriented approach required labeling the areas of KSAs that contributed to successful Academy performance. Detailed knowledge of the tasks allowed one to specify the KSA requirements for effective criterion performance (Fleishman, 1975). With respect to Fleishman's (1975) linking of abilities to task requirements, Landy (1989) discussed a job analysis conducted for fire fighters. He made a list of the most important cognitive abilities required to perform well on the job (Landy, 1989; Schmitt & Landy, 1993). Ekstrom (cited in Crosby, Rosenfeld & Thornton, 1979) evaluated similar cognitive domains for police officers.

Out of the sets of cognitive abilities obtained from previous job analyses, the test developers found that eight of those abilities applied directly to the UTSP officer domain. They were as follows: deductive reasoning, inductive reasoning, information ordering, memorization, problem sensitivity, spatial orientation, verbal comprehension, and verbal expression (Crosby et al., 1979; Landy, 1989). The following list provides definitions of the eight cognitive abilities related to police performance (Crosby et al., 1979; Fleishman, 1975; Landy, 1989):

1. Deductive reasoning refers to the ability to apply general rules to specific situations. It also refers to the

ability to proceed from stated principles to sensible conclusions.

2. Inductive reasoning implies the ability to find a rule or law which applies to a given situation. It includes determining a sensible explanation for a series of events that appear unrelated.

3. Information or semantic ordering is the skill involved in applying rules to a situation in order to align information in the most meaningful, intelligent sequence.

4. Memorization implies the ability to memorize and retain new facts that occur as a routine feature of a task or a job.

5. Problem sensitivity refers to the ability to detect a problem as a whole in addition to the problem's elements. It also refers to an understanding of the consequences associated with specific situations.

6. Spatial orientation is the ability to keep a clear idea of where one is in relation to the area or space that one occupies. It is also the ability to comprehend the position of a visual object in space.

7. Verbal comprehension expresses the ability to understand words and ideas, whether they are written or spoken.

8. Verbal expression refers to the ability to use either oral or written language to communicate information, ideas, and thoughts to others. This ability includes grammar, syntax, and vocabulary skills.

The job analysis has identified those KSAs most important to assess with a written entry-level test. The eight cognitive abilities listed apply particularly to the requirements of a new police officer. It is assumed that recruits attain additional skills and abilities through experience on the job. Because most, if not all, of the job applicants require training to do the job, there is no better predictor of training performance or on-the-job performance than general cognitive ability (Hunter & Schmidt, 1989).

The researchers requested opinions regarding the importance of the eight cognitive abilities from the UTSP departments. The responses reflected over 900 years of police experience. In descending order of importance, the departments rank-ordered the cognitive abilities as follows: Deductive Reasoning, Verbal Comprehension, Information Ordering, Verbal Expression, Inductive Reasoning, Memorization, Problem Sensitivity, and Spatial Orientation.

The purpose of personnel selection tests should be to determine the common ability requirements of different tasks (Blum, 1964; Schmitt & Landy, 1993). All responses indicated

that each of the eight cognitive abilities listed were, in fact, common requirements of most entry-level police tasks. Selecting such job-related abilities for testing was crucial because if the tested abilities were not truly required on the job, then the test told nothing about the applicant's prospective merit as an employee (Cronbach, 1970). Because each cognitive ability related to successful job performance, each demanded inclusion on the test.

To develop a preliminary test, the researchers studied job analysis data. They followed content-oriented procedures and drew questions from constructs linked to police performance. In addition, they reviewed many published and unpublished validation studies of police officer selection and other types of employment selection.

The ability-oriented job analysis facilitated the development of an ability test. According to Schmidt (1993), measures of ability "are among the most valid and useful predictors of performance in most jobs" (p. 498). In their review of validity generalization studies, Zeidner and Johnson (1991) found that for entry-level positions, "ability tests, in comparison with alternative predictors, are the best predictors of most types of job-performance criteria conventionally used" (p. 89).

Meta-analysis has pointed to the conclusion that ability tests are valid for predicting success in training and in actual job performance (Cook, 1988; Hunter, 1983; Zeidner & Johnson, 1991). Others also endorse the use of cognitive ability testing due to the tremendous quantity of evidence for positive non-zero validities of such tests across jobs (Hunter & Hunter, 1984; Sackett & Arvey, 1993).

An important question data needs to answer is why cognitive ability predicts job performance (Hunter & Schmidt, 1989). Data on job knowledge has indicated that cognitive ability predicts job performance because it is indicative of both learning capacity and job mastery potential. Hunter and Schmidt (1989) have shown that cognitive ability is highly correlated with job knowledge, and job knowledge is highly correlated with job performance. Through path analysis, they have found that an indirect causal path accounts for a large percentage of the effect of cognitive ability on job performance (Hunter & Schmidt, 1989).

Furthermore, meta-analysis demonstrated cognitive ability tests to be fair to minority applicants (Hunter & Schmidt, 1989). It is noteworthy that blacks usually score one to one-and-one-half standard deviations below whites on general cognitive tests (Arvey & Faley, 1988). Although blacks and Hispanics tend to have lower average scores and lower

hiring rates, the differences are not necessarily due to bias in the tests which predict performance (Schmidt, 1993). In comparison with other various selection devices such as assessment centers, Pynes and Bernardin (1989) supported the use of cognitive ability tests for entry-level police selection. For all of the reasons stated above, the preliminary test design included problems tapping all eight cognitive abilities which were associated with high performance on the job.

To facilitate sound test construction, Roznowski (1987) suggested that item content be selected extensively within the conceptual definition of the ability to be assessed. The test items fell into nine subtest areas. The definitions of the abilities to be assessed in nine subtest areas were as follows:

1. The ability to apply police information was the candidate's ability to apply training materials presented during the test. The training materials related directly to laws involved in police work.
2. The ability to perceive relationships referred to the candidate's ability to derive relationships between both concrete and abstract concepts. The subject matter of all concepts used in this test battery was police work.
3. The ability to compute basic work-related math problems was the candidate's ability to select and apply basic

mathematical functions and skills. Scenarios requiring computations were all associated with common police work.

4. The ability to comprehend written information was the candidate's ability to read and understand materials related to police work. In addition to comprehending the materials, the candidates had to draw inferences from them as well.

5. Grammar, spelling, and vocabulary referred to the candidate's language skills and familiarity with word meanings. All phrases were set contextually in police situations.

6. Recall ability was the candidate's ability to observe and recall important details and main ideas in two video segments portraying police scenarios.

7. Digit substitution ability was the candidate's ability to follow a code.

8. Spatial reasoning was the candidate's ability to see how disorganized pieces fit to make a whole.

9. The ability to convey thoughts clearly was the candidate's ability to write a summary of events in a clear, concise manner. The stimulus events were videos of two police situations.

It should be noted that nearly all portions of the test allowed the candidates to determine the relevance of the test items.

However, high performance did not require prior police experience or training.

The linkage between the cognitive abilities and the subtest areas helps in evaluating the degree to which the job content domain has been sampled (Sackett & Arvey, 1993). The linkage between the eight cognitive abilities and the subtest areas appears in Table 2.

Insert Table 2 about here

One suggestion the UTSP departments gave was to add a test section which would evaluate a candidate's common sense. The constructs of common sense and situational judgment were somewhat elusive and very difficult to assess through a written test. However, the test developers avoided using a large number of abstract situations while composing reasoning questions. It was their belief that depicting practical situations in items requiring reasoning would most clearly evaluate common sense.

The researchers determined that half of the eight critical cognitive abilities could be completely appraised through standard "question and answer" techniques. Assessing information ordering, memorization, verbal comprehension, and

verbal expression required innovative methods of presenting testing materials. The rationale for them appears below.

According to IPMA's (1992) study, two of the most important contributors to police success were the ability to learn proper police procedures and the ability to learn laws to be enforced. To assess the two abilities, IPMA developed a study booklet. Candidates studied its contents and responded to related questions. This method enabled the test administrators to maintain security of study materials and to increase testing standardization.

Thorndike (1962) mentioned some advantages of using motion pictures to present testing material. The advantages included the following: standardization/uniformity, objectivity, time control, and realism (Thorndike, 1962). The researchers in this study opted for a video instead of a booklet. Test administrators were able to maintain control over the presentation of the materials, to preserve test security, and to provide consistent testing atmospheres.

Job Similarity Determination.

The third purpose for job analysis data is to ascertain job similarity (Sackett & Arvey, 1993). In the case of the UTSP, this step is needed to determine whether jobs among campuses are similar enough to justify a common testing procedure. The researchers needed to determine how similar

the typical police officer's job was from one campus to another. The rationale was that if the fundamental requirements for success as a police officer were very similar from one location to another, then a justification existed for a common testing procedure.

The target occupation under analysis was that of patrol officer since new officers routinely begin their careers in this domain of work. Functions requiring additional training or experience such as those of an inspector were not of primary interest. The police departments at each campus work for the whole UT System and train at the UTSP Academy in Austin. Although different environments and communities surround each campus, the training each candidate receives at the Academy is consistent. Because all recruits train to be entry-level law enforcement officers on a college campus, one can correctly assume their jobs are equivalent. It is also a fact that UTSP officers can transfer from one campus to another. Therefore, it is logical to search for identical abilities among recruits from all campuses.

Test Validation

Once the entry-level test has been developed, the validation process begins. According to the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on

Measurement in Education (NCME) (1985), validity refers to the appropriateness and usefulness of the inferences drawn from performance on a given test. The Uniform Guidelines (EEOC, 1978) require that the validity of a job qualification test be proved by empirical data. The data must demonstrate that the test is predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated. The Standards for Educational and Psychological Testing (AERA, APA, & NCME, 1985) and the Uniform Guidelines (EEOC, 1978) also specify that one of the following types of validation methods must be used: construct validity, content validity, or criterion-related validity.

A construct validity study consists of the demonstration that a selection procedure measures candidates' identifiable characteristics which are important for successful performance on the job. A content validity study shows that the content of a selection device consistently represents the important aspects of performance on the job. A criterion-related validity study indicates the degree to which a selection test is predictive of or significantly correlated with important aspects of job performance. In addition to satisfying the requirements of the legal and ethical guidelines, a selection device should also have the following qualities:

practicality in application, ease of administration, fairness in design, and acceptance by candidates (IPMA, 1992).

While developing the test, the researchers stressed both concurrent and predictive validity in their study. They adhered to both ethical and legal validity guidelines, emphasizing criterion-related validity. Criterion-related validity studies require the correlation of a predictor (such as the entry-level police officer test) with a criterion (a measure of performance). "The main limitation to using criterion-related validity within the prediction context lies in the limited adequacy of the available criterion measures" (Thorndike, Cunningham, Thorndike, and Hagen, 1991, p. 140).

Criteria Selection.

There are four main qualities desired in a criterion measure. Thorndike et al. (1991) have listed them in descending order of importance as follows: relevance, freedom from bias, reliability, and availability. The relevance of a criterion measure refers to its correspondence to success on the job. Consistency of this correspondence implies reliability. Freedom from bias ensures that the correspondence to success on the job relates to actual competence and not personal characteristics. Finally, availability refers to the practicality and convenience associated with obtaining the selected measure.

The fundamental problem in selection research is the procurement of satisfactory criterion measures of performance on the job (AERA, APA, & NCME, 1985; Cronbach, 1970; Thorndike et al., 1991). Hirsh, Northrop, and Schmidt (1986) have found that it is difficult to develop a good criterion measure of job performance for police officers. This is due to the fact that police officers often work alone or work exclusively with a partner. It is possible that the few behaviors supervisors observe are unrepresentative samples. This does not allow for adequate observations by supervisors and impedes objective performance evaluations.

Performance evaluations may be subject to influences beyond an employee's control. Typical rating problems include generosity error, low reliability, the halo effect, and questionable validity (Thorndike et al., 1991). When evaluators are unwilling to give subordinates low ratings, the rating scores generally collect at the high end of the rating scale. This problem has been labeled the "generosity error." While the error itself produces limited differentiation among individuals, the inconsistent degree of error among raters is even more troublesome. Corrections for differences among raters are extremely difficult to implement (Thorndike et al., 1991). The significant between-rater discrepancies result in low reliability for the rating procedure.

Performance evaluators are often subject to the "halo effect" or the "pitchfork effect" in which they project a single strength or weakness into general competence or incompetence (Dobrish, Wolff, & Zevnik, 1984; Feltham, 1988). Some raters allow personality traits to contaminate performance judgments even when the only information available to them is strictly behavioral (Krzystofiak, Cardy, & Newman, 1988). When a rater allows irrelevant considerations and impressions to influence evaluations, the performance assessment becomes invalid (Froemel, 1979).

Raters also have a tendency to rank individuals similarly across different dimensions of job performance even when the dimensions are well-differentiated (Feltham, 1988). Ratings are often erratic and can be affected by many factors other than performance (Thorndike et al., 1991). Subjective, arbitrary judgments hinder performance ratings from being useful criterion measures. Schmitt, Gooding, Noe, and Kirsh (1984) also advise against the use of performance ratings. They note that performance ratings usually yield lower levels of validity than do more objective criteria (Schmitt et al., 1984).

To combat problems posed by rater subjectivity, it has been suggested that academic grades relevant to job performance are likely criteria (Thorndike, 1962; Thorndike et

al., 1991). Many have found that the type of job knowledge acquired in training programs, such as the UTSP Academy, plays a strong causal role in determining later job performance capabilities (Schmidt, Hunter & Outerbridge, 1986; Spielberger, Ward, & Spaulding, 1979). In both Abbatiello's and Shealy's papers presented at annual conventions of the American Psychological Association (cited in Spielberger et al., 1979), it has been demonstrated that measures of ability and intelligence are usually good predictors of academic performance specifically at police academies. For the above reasons, the criteria that have been selected for the purposes of this study are the academic averages and the final averages of the officers' police academy performance.

The researchers' final decision to correlate test performance with police academy performance was based upon the outcome of Washington v. Davis (1976). In this case the Supreme Court essentially ruled that academic performance in the police academy was an acceptable criterion measure in validity studies (Arvey & Faley, 1988).

Many would agree with Russell (1984) that "the superior strategy is to validate training with job performance" (p. 266). However, in many organizations it has been demonstrated that accurate and objective job performance ratings are

unavailable. In adhering to recommendations made in the Uniform Guidelines (EEOC, 1978) and the testing principles dictated by the AERA, APA, and NCME (1985), training performance should only be substituted for job performance when training performance is job-related (Arvey & Faley, 1990). The UTSP Academy strictly provides job-related training; therefore, the choice to use unbiased police academy performance records as criteria is justified. Criterion scores were available at the same time as scores on the predictor. The relevance, freedom from bias, reliability, and availability of the Academy performance records also justify their use as criteria.

Method

Subjects

Over a nine month period from 1992-1993, the UTSP participated in the development and validation of an entry-level police examination. Two conflicting goals affected the selection of test subjects. The first goal was to administer the test to as many UTSP officers as possible because larger samples generally provide more dependable correlations than do smaller samples (Cronbach, 1970). At the same time, the second goal was to test a sample that closely resembled an entry-level group or a sample with short tenure. Officers with less seniority obviously resembled an applicant pool better than seasoned officers with more seniority. The rationale was that by limiting the sample to officers with little or no seniority, the test results would not be skewed by long-term experience on the job. Minimizing the effect of job experience facilitated the researchers in generalizing results to entry-level candidates.

The researchers reconciled the conflicting goals by testing every available police officer with as few years on the job as possible. Although most of the subjects had between one and five years of experience as police officers, some sergeants and lieutenants took the test. The supervisors who took the test were retained in the study. Separate analysis of

their scores indicated that their inclusion did not bias the test results in any way.

A total of 197 police officers currently employed by the UTSP were tested during the course of the study. The entire sample consisted of 38 police cadets at the UTSP Academy and 159 job incumbents.

Cadet Sample.

Testing cadets in the UTSP Academy was beneficial in determining the predictive validity of the entry-level test. The cadets' test (predictor) scores were available before their criteria scores. Therefore, their test scores were used to predict their performance at the Academy. Determining such predictive validity was extremely important for ethical, legal, and practical reasons.

For admission to the UTSP Academy, candidates need a high school diploma (or equivalent), a passing score on an entry-level selection test, and a passing score on an agility test. Candidates must meet one of the following requirements: (a) 60 hours of college credit, (b) 30 hours of college credit and one year as a police guard, (c) two years as a police guard, or (d) two years of active military service plus one year as a police guard. Job candidates are also medically and psychologically evaluated. In future classes, candidates will be strictly required to have 60 hours of college credit and will

not be subjected to psychological evaluations. On the basis of these procedures, the decision is made to accept or to reject candidates for entry to the UTSP Academy.

Thorndike (1962) suggested that a preliminary form of a selection test should be administered to a sample of the same population with which the test would be used in final form. The UTSP planned to use the final form of the test in future selection procedures for the Academy. For these reasons, 38 officers comprising two Academy classes took the test. There were 17 cadets in the first (September - December 1992) class and 21 cadets in the second (February - May 1993) class.

There were cadets in the Academy classes who had been sent by the Texas Department of Public Safety (DPS). The DPS cadets were selected from a different applicant pool than those hired by the UTSP. DPS has different hiring procedures and recruiting processes, both of which may influence the composition of applicant pools. Although DPS cadets took the entry-level test, their scores were not combined with UTSP cadet data. The scores of the DPS cadets were studied, however, to aid in test development. By separating the UTSP cadets from the DPS cadets, the results of the study could more appropriately be generalized to the applicant pool for UTSP positions.

The race and gender compositions of the cadet sample were as follows: black females = 1; black males = 5; Hispanic females = 1; Hispanic males = 6; white females = 6; white males = 19.

Incumbent Sample.

Due to budget and personnel constraints, not all incumbents were able to take the entry-level examination. The list of campus police departments that provided incumbents for testing appears in Table 3.

Insert Table 3 about here

A total of 159 job incumbents took the test. For the incumbents, criteria scores were achieved prior to predictor scores. Therefore, the incumbents provided concurrent validity measures for the entry-level test.

The race and sex compositions of the incumbent sample were as follows: black females = 7; black males = 22; Hispanic females = 4; Hispanic males = 19; white females = 18; white males = 89. The comparison among compositions of the cadet sample, the incumbent sample, and the entire UTSP force appears on Table 4.

Insert Table 4 about here

Procedure

To ensure testing environment consistency, the same experimenter administered the test to all subjects. Test materials were stored in a locked office, and individual test results were not available to any non-participating members of the UTSP. Neither the training instructors at the Academy nor the police supervisors were aware of the officers' entry-level test performance. Therefore, all subjects were assured that performance on the entry-level test had no bearing on hiring decisions or on performance evaluations. Only after completion of data analysis were subjects allowed to see their scores on the test. All subjects received confidential, individual feedback. Since the data was not released until all subjects took the test, the test measures remained uncontaminated.

First Cadet Sample.

As mentioned, the entry-level test tapped all eight cognitive abilities deemed vital to police performance. The test was divided into four sections. For the first cadet sample ($n = 17$), the test consisted of the following four sections:

1. Cadets took the main body of the test which consisted of 74 multiple-choice questions. This segment was referred to as the Main Test. Within the Main Test were 10 vocabulary test items taken from Alwin (1988). This subsection was referred to as Vocabulary. Additional analysis was performed on the Main Test without Vocabulary to assess the meaningfulness of the Vocabulary scores.

2. Cadets viewed a video of actual footage from a traffic stop that went awry in Refugio, Texas. After the video, candidates answered 12 questions related to the incident which they had just witnessed. This segment was referred to as the Refugio Test.

3. Cadets had four minutes to complete the Spatial Reasoning Task, taken from the Revised Beta Examination (Kellogg & Morton, 1962).

4. Cadets viewed a second video. The video portrayed two police officers responding to a domestic disturbance call. Candidates had eight minutes in which to write all pertinent details about the situation which they had just witnessed. This task was selected primarily because job analysis indicated the existence of a large emphasis on writing reports, following instructions, and testifying in court.

First Test Shakedown.

After the first cadet group was tested, item difficulty was assessed. The researchers considered subject feedback and basic statistics in order to “shake down” the test (IPMA, 1992). This process helped eliminate major flaws before the second cadet class and job incumbents took the test. In assessing item difficulty, the smaller the percentage correctly answering a given item, the more difficult the item (Thorndike, 1962). In the test shakedown, all items deemed too easy or too hard were changed or removed from the test because they failed to appropriately discriminate between the more and less capable candidates.

The rational judgment, expertise, and opinions of the cadets were welcomed. They commented on the length of the test, the difficulty of the questions, the language used, and the perceived fairness to various race and gender groups. The feedback was encouraging, implicating high face validity. Each cadet who commented believed the test was a marked improvement over the entry-level test they had taken during their actual selection process. They believed the difficulty level of the items was reasonable. One officer pointed out that one of the questions regarding Satanic worship symbols had racial prejudice implications. Such questions evoking sensitivity were marked for either dismissal or alteration.

Due to the small sample size of 17 subjects, results were interpreted merely as general indicators of areas requiring improvement rather than as effective and reliable predictors of future results. Based on cadet comments and the item analysis data, several item revisions, additions, and deletions were made. Directions were clarified; distractors were made more difficult or more easy depending on the data; and item groups were organized in a more logical sequence.

Subsequent Samples.

Analysis of the first cadet sample test results indicated the need for additional basic mathematics and inductive reasoning items. The researchers also decided to add an additional section taken from the Revised Beta Examination (Kellogg & Morton, 1962) (see Appendix A for instructions). The second cadet sample ($n = 21$) and the incumbent sample ($n = 159$) took a test that consisted of five sections. The five sections were organized as follows:

1. Subjects took the Main II Test, which consisted of 100 test items (see Appendix B for Main II Test items). The Main II Test was a revision of the Main Test that the first cadet sample took. To revise the Main Test, four items were removed, leaving 70 of the original 74 test items. Thirty new items were added to comprise the Main II Test. Additional

analysis was performed on the Main II Test without Vocabulary to assess the meaningfulness of the Vocabulary scores.

2. Subjects had two minutes to complete the Digit Substitution Task, taken from the Revised Beta Examination (Kellogg & Morton, 1962).

3. Subjects viewed the Refugio, Texas video and took the Refugio Test (see Appendix C for Refugio Test items).

4. Subjects had four minutes to complete the Spatial Reasoning Task, taken from the Revised Beta Examination (Kellogg & Morton, 1962).

5. Subjects viewed the domestic disturbance video and wrote all pertinent details about that which they had just witnessed (see Appendix D for instructions).

Second Test Shakedown.

The second cadet class and the job incumbents also responded favorably to the entry-level examination. Several officers mentioned that the test items requiring the application of laws were up-to-date with the current Texas Criminal Laws. The officers noted the test topics and individual test questions maintained a high degree of job-relatedness. The more forms of evidence there are available to uphold a claim of job-relatedness, the better (Sackett & Arvey, 1993). The apparent perception of a high level of job-relatedness perhaps explained why the officers believed the

test was the fairest and the best of all tests they had ever taken. These comments implicated construct validity.

Data Entry and Analysis.

For the Main/Main II Test and the Refugio Test the candidates marked their answers on a score sheet. The test administrator examined all answer sheets to ensure they were properly completed. An optical scanner then read the scores into a computer file. Test response analysis software detected additional errors in the answer sheets such as cloudy erasures, missing social security numbers, skipped answers, and multiple responses. Answer sheets were corrected and re-entered accordingly.

The Spatial Reasoning and Digit Substitution Tasks were manually scored by following the grading instructions in the Revised Beta Examination Manual (Kellogg & Morton, 1962).

The writing samples describing the domestic disturbance video were transcribed into a word processing program (Microsoft Word 5.1a) on a Macintosh Quadra 700. All grammar, punctuation, and spelling errors were not corrected during transcription. The program provided an objective measurement of the reading ease and writing (grade) level of the writing samples. All scoring methods were available and feasible for consistent implementation at all UTSP institutions.

After test data was obtained, the researchers gathered UTSP Academy performance data. For more recent classes, the data was in computer files. For less recent classes, the UTSP departments voluntarily gave information regarding individuals' UTSP Academy performance. Academic averages and final averages were calculated and collected for all officers.

Each officer attained an academic average. This was the average of the scores achieved on the major tests taken during the Academy. Each week the cadets were tested on the topics covered in academics. The weekly test topics were as follows: (a) Penal Code and Texas Motor Vehicle Laws; (b) Family Code and Education Code; (c) Code of Criminal Procedure, Probable Cause, and Collision Investigation; (d) Driving While Intoxicated (DWI), Patrol Techniques, and Crime Prevention; (e) First Aid, Report Writing, and Miranda Warning; (f) Note-taking and Narcotics; (g) Detection and Frisk, Search and Seizure, and Use of Force; (h) Handcuffing and Armed and Dangerous Criminals; (i) Firearms, Baton Usage, Defensive Tactics, Outside Agencies, and Driving Skills; (j) Crime Scene, Physical Evidence, Rules of Evidence, Confessions, Interrogations, Alcohol and Beverage Code, and Executive Protection; (k) Civil Process, Radio Communication, Bombs, Multiculturalism, and History of Policing; and (l) Crowd and Riot Control, Hazardous

Materials, Professionalism and Ethics, Problem Solving, and Critical Thinking.

The final averages consisted of a combined performance percentage. Fifty percent of the final average was the major exam average. In addition to the written exams, candidates took several proficiency skill exams including a firearms exam. Twenty-five percent of the final average was the Practical Problems Skill Test average. The tests assessed the candidates' proficiency in the following areas: (a) handcuffing, (b) night firing, (c) report writing, (d) collision reports and investigation, (e) crime scene protection and recording, (f) arrest, search, and seizure, and (g) subject interviewing and note taking. Ten percent of the final average was a physical training score. Another ten percent was the pop quiz average. The final five percent of the final average was a notebook score.

Results

Table 5 provides the means and variances of the

Insert Table 5 about here

independent and dependent variables from both the cadet sample ($n = 38$) and the incumbent sample ($n = 159$). Table 5 also presents the correlations between the variables.

Correlations between the entry-level test scores (predictors) and the criteria are called validity coefficients. For all predictors except the reading ease of the writing sample, the correlations with the criteria are positive.

It is noteworthy that validity coefficients rarely rise above 0.60 (Cronbach, 1970). Although 0.60 is far from perfect prediction, a positive correlation indicates that predictions from the test are more accurate than guesses (Cronbach, 1970).

Differences Between the Samples

The correlations between the criteria and the predictors were primarily higher in the cadet sample than in the incumbent sample. The largest difference (0.30) occurred in the correlation between the digit substitution task and the academic average. Generally, the differences between the

correlations for the incumbents and for the cadets were not as great.

Differences Between the Dependent Variables.

The correlations between the academic average and the predictors were consistently higher than the correlations obtained by using the final average as a criterion. The correlation strength trends in the final average correlations generally mirrored those in the academic average correlations. The main area of difference between the criteria lay in the magnitude of their correlation with the predictors.

If the variances of the dependent variables differed by a factor of three or more, then further analysis would be required to explain the disparity. As Table 5 demonstrates, the means and the variances of the dependent variables were very similar. The variances did not even differ by a factor of one. Therefore, no further analysis was required to explain the differences between the dependent variables.

Differences Among the Independent Variables.

Within any given correlation, the mean and variance of the independent variable involved was similar for both testing groups. When comparing the means and variances of the independent variables from the cadet group to those from the incumbent group, Table 5 demonstrated that the differences rarely exceeded a factor of two and never exceeded a factor of

three. Therefore, no further analysis was required to explain the differences found between the independent variables.

Statistics for the Main II Test and the Refugio Test

Table 6 presents data from the incumbent sample.

Insert Table 6 about here

Included in Table 6 are the following statistics for both the Main II Test and the Refugio Test: mean, standard deviation, and test reliability. The LXR Scoring program provided the test statistics. The reliability measure is the Cronbach alpha reliability coefficient. The coefficient alpha tells how well scores obtained under one condition (such as one testing period) represent universe scores (Cronbach, 1970).

The statistics were calculated for both the cadet sample and the incumbent sample. The statistics were also calculated for the race and gender groups within the incumbent sample: black, female, Hispanic, male, and white.

Main II Test.

On the Main II Test, the cadets averaged 81.57 items correct out of a possible 100. The incumbents' average of 81.03 was very similar to that achieved by the cadets. The overall Cronbach alpha reliability for the test was 0.89 for the cadets and 0.90 for the incumbents.

The white sample performed best on this test with a mean score of 84.02. The females' (82.03) performance was higher than the males' (80.71). The black sample (75.76) was followed by the Hispanic sample (74.41). The reliability was high among the different race and gender groups, ranging from 0.81 (female sample) to 0.91 (male sample).

Refugio Test.

On the Refugio Test, the cadets averaged 10.11 items correct out of a possible 12. The incumbents' average was slightly lower (9.71). The overall Cronbach alpha reliability for the test was 0.38 for the cadets and 0.59 for the incumbents.

The white sample performed best on this test with a mean score of 10.06. Males and females performed equally (9.72). The black sample (9.35) was followed by the Hispanic sample (8.97). The reliability was not as high as it was for the Main II Test. The reliability measures ranged from 0.44 (female sample) to 0.65 (Hispanic sample).

Race and Gender Differences

Table 7 presents the correlations of the criteria with the

Insert Table 7 about here

predictors for the incumbents ($n = 159$). In order to assess adverse impact and fairness on the test, correlations were calculated separately for blacks ($n = 29$), females ($n = 29$), Hispanics ($n = 23$), males ($n = 130$), and whites ($n = 107$).

As Table 7 depicts, the correlations between the predictors and criteria (academic average/final average) varied among the groups. For instance, the digit substitution task was a better predictor for Hispanics (0.27/0.18) and males (0.22/0.20) than it was for the other groups. The Main II Test was a powerful predictor for all incumbents (0.59/0.57), especially for Hispanics (0.57/0.52), males (0.64/0.62), and whites (0.64/0.62). The Refugio Test was also a strong predictor for all incumbents (0.46/0.43). Specifically, the Refugio Test was a good predictor for Hispanics (0.47/0.43), males (0.51/0.47), and whites (0.43/0.36).

Reading ease was not a good predictor of performance on the criteria for any of the groups. All of the correlations involving reading ease were negative, ranging from -0.18 (Hispanic sample, academic average) to -0.04 (female sample, final average). The spatial reasoning task correlated positively with the criteria. Spatial ability was a better predictor with the black (0.20/0.16), Hispanic (0.21/0.26), and male (0.33/0.27) samples than it was with the other groups. Vocabulary correlated positively and strongly with the

incumbent sample (0.55/0.53), especially Hispanics (0.47/0.45), males (0.61/0.59), and whites (0.56/0.52). The writing level also correlated positively with the criteria. Writing level was most positively correlated with the criteria in the female sample (0.26/0.21).

For the entire incumbent sample, the four strongest predictors of performance on the criteria were the Main II Test, the Refugio Test, Spatial Reasoning, and Vocabulary. The best predictor of performance on the criteria (academic average/final average) for each group is as follows: all incumbents, Main II Test (0.59/0.57); black incumbents, Refugio Test (0.35/0.36); female incumbents, Main II Test (0.33/0.23); Hispanic incumbents, Main II Test (0.57/0.52); male incumbents, Main II Test (0.64/0.62); and white incumbents, Main II Test (0.60/0.53).

Test Adverse Impact and Fairness.

Table 8 specifically compares statistics from the Main II

Insert Table 8 about here

Test for different race and gender groups within the incumbent sample. The performance of the cadets was also compared to the performance of the incumbents in Table 8.

Although the disparity between group performance on the Main II Test appeared large, the difference between mean scores for black and white participants was not significant at the .01 level. The difference between mean scores for Hispanic and white incumbents was also not significant at the .01 level.

Male and female scores on the entry-level test were similar, with females scoring 1.32 points higher. There was not a significant difference at the .01 level in their performances. Cadets and incumbents, too, scored similarly on the Main II Test, with cadets scoring 0.54 points higher. There was not a significant difference between their mean scores at the .01 level

Attrition

To better understand the test data, the researchers collected additional data from the UTSP personnel database. Officers who graduated from the UTSP Academy promised to serve for at least two years. After two years, they were allowed to pursue other job leads.

The researchers hypothesized that attrition of high-performing minorities was reflected in the overall performance of the minority groups. The data collected included the performance averages of those officers attending the UTSP Academy during the last two years (those still bound

to the service contract). Those averages were compared to those of the officers who graduated from the Academy two to five years ago who are still employed by the UTSP.

Table 9 presents the attrition summary for Academy

Insert Table 9 about here

Classes 46 through 55. In Table 9 one can see that overall, the final averages of the resignees are greater than those attained by the classes as a whole. This implies that those who perform better in the Academy are more likely to resign, presumably because of better job offers outside the UT System.

Table 10 provides specific resignee data. For each race

Insert Table 10 about here

and gender group one can determine the following percentages: the percentage of total cadets who have resigned, the percentage of each race and gender group which comprise the resignee group, and the percentage of each race and gender group which comprise the entire UTSP force.

In summary, over the past five years, 42 officers out of the 197 Academy graduates (21.32 %) who have been eligible to

resign have done so. The only officer group that has a greater resignee percentage than UTSP composition percentage is the Hispanic group. All of the Hispanic resignees have been male. The percentage of Hispanic male officers who have resigned is 19.05, which exceeds the Hispanic male composition of the UTSP force (11.36 %).

Discussion

This study assessed both the concurrent and the predictive validity of an entry-level examination. Both strategies were appropriate for this criterion-related validity study. The majority of subjects in the study were incumbents; therefore, most of the data related to concurrent validation.

Concurrent validation procedures are less time-consuming and expensive than predictive validation procedures. Therefore, concurrent procedures are the most commonly used in gathering evidence for criterion-related validity studies (Arvey & Faley, 1988). However, simultaneously gathering test and job information from current employees and correlating the two sets of scores can limit the ability to generalize test results.

Criticisms of Concurrent Validation Procedures

There are three main criticisms of concurrent validity studies. The first criticism is that test scores obtained from incumbents may not very well apply to others even though the researchers attempted to find a sample most representative of the applicant pool. Arvey and Faley (1988) suggest that incumbents have already found a job that fits their particular skills and personality whereas applicants are still seeking a good fit between a job and their needs. Therefore, the test

score range for incumbents is apt to be more homogeneous thus restricted than that for applicants.

It is also important to note that the incumbent group tested is not a random sample. In order to have a measure of job performance, the researchers are restricted to collecting performance data of cadets and of incumbents. However, the population of interest in employee selection is the applicant pool. For many reasons, incumbents have a much narrower range of cognitive ability than do job candidates (Hunter & Schmidt, 1989). Cognitive ability itself has a coefficient of generalizability of about 0.80, which further limits the potential of the study validity (Cook, 1988; Hunter & Schmidt, 1989).

There are several reasons for which incumbents have a much narrower range of cognitive ability than do applicants. In order to be hired by the UTSP, the officers have obviously done reasonably well on other tests in the past. Compared to the estimated test performance of an applicant pool, fewer low scores in an incumbent sample are expected. It is important to realize that differences between applicants and incumbents are likely to be even greater than those between cadets and incumbents. It is highly probable that the entry-level test will produce a much larger standard deviation when it is implemented with actual candidate groups. The

researchers recommend that future studies involve the collection of information from job applicants and compare their test scores with those of job incumbents to detect differences between the groups.

Another problem presented by concurrent validation design is that incumbents may change over time by developing those skills required for success on the job (Arvey & Faley, 1988). For instance, police officers must refresh their basic mathematical skills in order to fill out certain forms such as mileage logs. Job applicants may have let such skills lie dormant prior to employment.

The third criticism of concurrent validity studies is that job incumbents may respond to tests differently than actual job applicants (Arvey & Faley, 1988). It is desirable to perform validation testing under conditions of motivation and experience similar to those likely to be encountered in practical use (Thorndike, 1962). However, incumbents may be less likely to give their greatest effort on an entry-level test since their test performance has no effect on their careers. It is possible that those incumbents who have seen little personal advantage in taking the entry-level test may have yielded meaningless scores.

Even though the cadets in this study have been hired only recently, there is still a difference between their mental set

and the mental set of an actual applicant. While performing future studies, it may be helpful to encourage incumbents to adopt job applicant motivation and to do as well as they can on the test.

Test Evaluation

Three main questions must be addressed in deciding the appropriateness of a test for the purpose of selecting cadets for the UTSP Academy. The first question asks if the test does, in fact, predict performance. The second question asks what conditions must be met in order for the test to predict performance. The third question asks if the test yields similar results for different groups.

The answer to the first question, whether or not the test predicts performance, is yes. The validity of the test has been supported both in terms of the methods used and of the adherence to legal guidelines. It has also been supported by the relationship between test scores and UTSP Academy performance. As the results demonstrate in Table 5, the correlations between predictors and criteria are higher than chance would suggest. With the exception of reading ease, all predictors positively correlate with performance at the UTSP Academy.

Even though the entry-level test has been demonstrated to successfully predict performance, there are conditions that

must be met in order for the test to serve its purpose. Sample size, restriction of range, and an organization's selection ratio can have significant impact upon the results of a criterion-related validity study.

Sample Size.

It is highly probable that both the total number of cadets and of incumbents participating in the study have affected the results of the study. Although the entire UTSP force employs almost 300 officers, the minority compositions are small. When an insufficient number of incumbents is available, it is technically difficult to perform a criterion-related validity study. This limits the confidence of the criterion reliability estimation (Sackett & Arvey, 1993). There is a great need for criterion-related research in the field of industrial psychology (Schmitt & Robertson, 1990). Therefore, despite small sample sizes, researchers should strive to perform such research.

The differences calculated between minority and non-minority performance were smaller than those normally yielded by cognitive ability tests. However, a prediction of a candidate group's majority-minority difference should not be made from these data due to the sample size. Although the validation study captured a large portion of the UTSP minority population, the UTSP minority compositions themselves are small.

Restriction of Range.

There is probably restriction in test score range in both the cadet and incumbent samples. The UTSP draws applicants from a relatively narrow range of ability. This fact is due to a limited recruiting budget and a limited salary fund. Because the UTSP is a narrow-range group, validity coefficients are bound to be lower than they would be in a wide-range group (Cronbach, 1970). To become a wide-range group, the UTSP would require additional recruiting funds, more extensive recruiting programs, and increased salary offers to attract applicants of increased quantity and higher quality.

Attrition may cause restriction of range on the criteria. As Table 9 has demonstrated, the final averages of the resignees are greater than those attained by the classes over the past five years. One may infer from this that those who perform well at the Academy may have a greater opportunity to find other, perhaps higher-paying, police work outside the UTSP. The attrition of the higher performers, especially those who belong to minority groups, further decreases the variation found among the incumbents.

Selection Ratio.

The value of a correlation heavily depends on selecting fewer people than apply. The relationship between the number of people who apply and the number of people who are hired is

the selection ratio. As the Taylor-Russel Tables (as cited in Landy, 1989) demonstrate, the utility of any selection procedure varies dramatically with the selection ratio. If an organization hires nearly all applicants, then an entry-level test is hindered from performing its function of selecting. Therefore, an organization must attract as many qualified applicants as possible in order to fully utilize the selection capabilities of a sound entry-level test.

Because the effects of sample size, restriction of range, and the selection ratio can enhance or decrease the relationship between predictors and criteria, continued research is needed to completely answer the second question.

Fairness to Minority Groups That Include Females.

The third question asks whether or not the test yields similar results for different groups. In essence, the question addresses fairness to minority groups that include females. Failure to use accurate predictors of job performance can cause tremendous productivity losses in the economic sector. However, usage of such tests also has the potential to minimize minority representation in the workforce. It is extremely difficult if not impossible to resolve the conflict between America's need to be internationally competitive and America's attempt to provide equality for minorities.

Although ability measures are among the most useful predictors of job performance, ability tests are often subject to lawsuits due to the adverse impact they have had on minorities (Arvey & Faley, 1988). Accumulated evidence shows that blacks and Hispanics have lower average scores and subsequently lower hiring rates (Guion, 1993). This should not be surprising since "*aggregate* differences between groups will likely occur on *any* reliably measured variable" (Arvey & Faley, 1988, p. 122). Such a claim is not meant to imply inferiority. It merely encourages additional research to assess the degree to which test differentials can predict job-relevant behavior.

The results of this study indicated that there were no significant differences in performance between the minority and majority groups. However, fewer than 30 incumbents from each minority group took the test. Due to the small number of minority incumbents tested, as well as the small number of minority incumbents employed by the UTSP, it would be unfair and inappropriate to conclude any answer with confidence. Arvey and Faley (1988) agree that when there are small samples involved, any conclusions about differential validity of males and females in the employment arena must wait for the collection of further data.

From the results presented in Table 7, it is apparent that different independent variables predict success better than others for the various groups. Therefore, it is important to continue testing the different predictors. Validation itself is a continuous process. More data can either confirm or reject the decision to include specific sections of the test in the future.

Recommendations

At this stage, the entry-level test is ready for implementation. The most current form of the test reflects all test shakedowns and alterations made due to item analysis and to participant comments. The only difference in grading the test will be the omission of the reading ease score. The only score to be taken from the writing sample will be writing level. Over time, the UTSP can evaluate the cost-effectiveness of this portion of the test. At this point, the writing score has the potential to be a valuable predictor of performance at the Academy. It was also the third highest predictor for performance in the female sample, and the female group is definitely a group of interest for continued validation.

The UTSP can give the new entry-level test for selection of the next two Academy classes. This will not only increase sample sizes of minorities, but it will also enable the UTSP to

include actual job applicants in the subject pool. Because the cadets and incumbents tested in this study were a preselected group, adding applicant scores to the data will introduce more variance. Once an ample number of applicants is tested and the minority sample sizes increase, then the UTSP can assess test fairness through a study of group differences in regression lines.

Additional studies on the predictive validity of selection tests used in hiring entry-level police officers need to be undertaken. Agencies must also remain vigilant and comply with changes in employment selection law. By attempting to improve its selection procedures, an organization significantly increases its legal leverage should a discrimination case arise. While the above analyses do not provide blanket endorsement for its use, the new entry-level test is a valid and fair predictor of UTSP performance. It has laid the foundation for the accurate and fair selection of competent entry-level police officers.

References

- Alwin, D. F. (1988). Cohort differences in vocabulary knowledge in the United States. Ann Arbor, MI: University of Michigan, Department of Sociology and Institute for Social Research.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). Standards for educational and psychological tests. Washington, DC: American Psychological Association.
- Arvey, R. D., & Faley, R. H. (1988). Fairness in selecting employees (2nd ed.). Reading, MA: Addison-Wesley.
- Blum, R. H. (Ed.). (1964). Police selection. Springfield, IL: Charles C. Thomas.
- Brown Maroney. (1991). Law watch. Austin, TX: Brown Maroney & Oaks Hartline.
- Cascio, W. F. (1987). Costing human resources: The financial impact of behavior in organizations (2nd ed.). Boston: PWS-Kent.
- Civil Rights Act of 1991, 42 U. S. C. 2002e-2 (1991).
- Colarelli, N. J., & Siegel, S. M. (1964). A method of police personnel selection. Journal of Criminal Law, Criminology, and Police Science, 55, 287-289.

Cook, M. (1988). Personnel selection and productivity.

New York: John Wiley & Sons.

Cronbach, L. J. (1970). Essentials of psychological testing

(3rd ed.). New York: Harper & Row.

Crosby, A., Rosenfeld, M., & Thornton, R. F. (1979).

Development of a written test. In C. D. Spielberger (Ed.),

Police selection and evaluation: Issues and techniques

(pp. 143-153). Washington: Hemisphere.

Dobrish, C., Wolff, R., & Zevnik, B. (1984). Hiring the right

person for the right job. New York: Franklin Watts.

Dunnette, M. D. (1966). Personnel selection and placement.

Belmont, CA: Wadsworth.

Equal Employment Opportunity Commission, Civil Service

Commission, Department of Labor, & Department of

Justice (1978). Uniform guidelines on employee

selection procedures, Federal Register, 43, 38295-

38309.

Fabricatore, J. M. (1979). Preentry assessment and training:

Performance evaluation of police officers. In C. D.

Spielberger (Ed.), Police selection and evaluation: Issues and

techniques (pp. 77-85). Washington: Hemisphere.

- Feild, H. S., Holley, W. H., & Buckner, K. E. (1990). Personnel selection in police and fire departments: A study of employment discrimination case characteristics and outcomes. Labor Law Journal, 9, 622-632.
- Feltham, R. (1988). Validity of an assessment centre: A 1-19-year follow-up. Journal of Occupational Psychology, 61, 129-144.
- Fleishman, E. A. (1975). Toward a taxonomy of human performance. American Psychologist, 30, 1127-1149.
- Froemel, E. C. (1979). Objective and subjective measures of performance. In C. D. Spielberger (Ed.), Police selection and evaluation: Issues and techniques (pp. 87-111). Washington: Hemisphere.
- Gatewood, R. D., & Feild, H. S. (1990). Human resource selection. Hinsdale, IL: The Dryden Press.
- Grider, K., & Bailey, T. A. (1992). The civil rights act of 1991: Congress reverses the supreme court. Texas Bar Journal, 1, 37-40.
- Guion, R. M. (1966). Employment tests and discriminatory hiring. Industrial Relations, 5, 20-30.
- Guion, R. M. (1993). The need for change: Six persistent themes. In N. Schmitt, W. C. Borman, & Associates (Eds.), Personnel selection in organizations (pp. 481-515). San Francisco: Jossey-Bass.

- Hirsh, H. R., Northrop, L. C., & Schmidt, F. L. (1986). Validity generalization results for law enforcement occupations. Personnel Psychology, 39, 399-420.
- Hunter, J. E. (1983). A causal analysis of cognitive ability, job performance, and supervisory ratings. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), Performance measurement and theory (pp. 257-266). Hillsdale, NJ: Erlbaum.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-98.
- Hunter, J. E., & Schmidt, F. L. (1989). Meta-analysis: Facts and theories. In M. Smith & I. T. Robertson (Eds.), Advances in selection and assessment (pp. 203-216). New York: John Wiley & Sons.
- International Personnel Management Association. (1988). Development and validation of the A-2 entry-level police officer examination. Unpublished manuscript.
- International Personnel Management Association. (1992). Development and validation of the A-3 police officer selection test for the international personnel management association (IPMA). Unpublished manuscript.
- Jurkanin, T. J. (1989). Job task analysis. FBI Law Enforcement Bulletin, 11, 10-13.

- Kellogg, C. E., & Morton, N. W. (1962). Revised beta examination. New York: Psychological Corporation.
- Kellogg, C. E., & Morton, N. W. (1962). Revised beta examination manual. New York: Psychological Corporation.
- Krzystofiak, F., Cardy, R., & Newman, J. (1988). Implicit personality and performance appraisal effectiveness. Journal of Applied Psychology, 73, 515-521.
- Landy, F. J. (1989). Psychology of work behavior (4th ed.). Pacific Grove, CA: Brooks/Cole.
- Martin, S. L., & Raju, N. S. (1992). Determining cutoff scores that optimize utility: A recognition of recruiting costs. Journal of Applied Psychology, 77(1), 15-23.
- Morrison & Foerster. (1992). Employment law commentary. San Francisco, CA: Author.
- Moynihan, J. B. (1992). Employment discrimination law in the federal sector: After the civil rights act of 1991. Texas Bar Journal, 5, 450-456.
- Pynes, J. E., & Bernardin, H. J. (1989). Predictive validity of an entry-level police officer assessment center. Journal of Applied Psychology, 74(5), 831-833.
- Roznowski, M. (1987). Use of test manifesting sex differences as measures of intelligence: Implications for measurement bias. Journal of Applied Psychology, 72, 480-483.

- Russell, J. S. (1984). A review of fair employment cases in the field of training. Personnel Psychology, 37, 261-276.
- Sackett, P. R., & Arvey, R. D. (1993). Selection in small n settings. In N. Schmitt, W. C. Borman, & Associates (Eds.), Personnel selection in organizations (pp. 418-447). San Francisco: Jossey-Bass.
- Sauls, J. G. (1992). The civil rights act of 1991: New challenges for employers. FBI Law Enforcement Bulletin, 9, 25-31.
- Schmidt, F. L. (1993). Personnel psychology at the cutting edge. In N. Schmitt, W. C. Borman, & Associates (Eds.), Personnel selection in organizations (pp. 497-515). San Francisco: Jossey-Bass.
- Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). The impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings on job performance. Journal of Applied Psychology, 71, 432-439.
- Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. Personnel Psychology, 37, 407-422.

- Schmitt, N., & Landy, F. J. (1993). The concept of validity. In N. Schmitt, W. C. Borman, & Associates (Eds.), Personnel selection in organizations (pp. 275-309). San Francisco: Jossey-Bass.
- Schmitt, N., & Robertson, I. (1990). Personnel selection. Annual Review of Psychology, 41, 289-319.
- Smart, B. D. (1983). Selection interviewing: A management psychologist's recommended approach. New York: Wiley.
- Spielberger, C. D., Ward, J. C., & Spaulding, H. C. (1979). A model for the selection of law enforcement officers. In C. D. Spielberger (Ed.), Police selection and evaluation: Issues and techniques (pp. 11-24). Washington: Hemisphere.
- Thorndike, R. L. (1962). Personnel selection: Test and measurement techniques. New York: John Wiley & Sons.
- Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen, E. P. (1991). Measurement and evaluation in psychology and education (4th ed.). New York: Macmillan.
- Washington v. Davis, 96 S. Ct. 2040 (1976).
- Zeidner, J., & Johnson, C. D. (1991). The economic benefits of predicting job performance: Selection utility (Vol. 1). New York: Praeger.

Appendix A

Instructions for Administering Subtests from the Revised Beta Examination Manual (Kellogg & Morton, 1962)

The instructions for administering the digit substitution task are as follows: Say, "Look at the divided row of boxes at the top of the page. In the top part of each box is a drawing. Underneath each drawing is a different number. Whenever you see the divided boxes in the three rows at the bottom of the page, I want you to put the right number under every drawing. Look at the first four boxes in the first row. They have been done, as an example, to show you how to do it. Notice that 3 has been put below pear, 1 below kitten, 2 below cross, and so on. Now you go ahead and do the rest. Put the right number under every mark. Go ahead." When everyone indicates mastery say, "Stop. Turn over the page. Put the right number every mark. Ready: Go!"

The instructions for administering the spatial reasoning task are as follows: Say, "Look at the examples at the top of the page. The directions say: 'Mark each square to show how the pieces at its left will fit into it.' The top three have been done to show you how to do it. You do the six at the bottom. Go ahead." When all have demonstrated an understanding of the instructions, say, "Stop. Turn over the page. Mark each square to show how the pieces at its left will fit into it. Ready: Go!"

Appendix B**Main II Test**

Instructions: There is only one correct answer for each question. When you select a response, mark the corresponding box on the answer sheet. Be sure that your shading is dark and neat.

(Note: When the test is compiled for printing, ensure that the items pertaining to one passage are arranged in such a way that the subjects do not have to unnecessarily flip the test pages back and forth.)

* Questions 1-3 are based on the following definitions. Do not assume anything to be true that is not stated in the definitions. Select the charge that fits the situation best.

- A person who, without the effective consent of the owner, enters a habitation, or a building (or any portion of a building) not then open to the public, with intent to commit a felony or theft commits burglary.

- A person who, by threatening, unlawfully appropriates property with intent to deprive the owner of property commits theft.

- A person who, in the course of committing theft, unintentionally, knowingly, or recklessly causes or threatens bodily injury to another, commits robbery.

- A person who commits robbery and uses or exhibits a deadly weapon commits aggravated robbery.

1. A man entered a neighbor's garage and took a mountain bike. According to the definitions above, the man should be charged with:

- A. burglary
- B. theft
- C. robbery
- D. aggravated robbery
- E. none of the above

2. A woman with a pistol approached a cashier and said, "Give me all the money in the register, or I'll blow you away." According to the definitions above, the woman committed:

- A. burglary
- B. theft
- C. robbery
- D. aggravated robbery
- E. none of the above

3. While stealing the woman's handbag, the man accidentally pulled her shoulder out of joint. According to the definitions above, the man should be charged with:

- A. burglary
- B. theft
- C. robbery
- D. aggravated robbery
- E. none of the above

* Questions 4-6 are based on the following definitions. Do not assume anything to be true that is not stated in the definitions. Choose the charge that fits the situation best.

- A person commits murder if he or she intends to cause serious bodily injury and commits an act clearly dangerous to human life that causes the death of an individual.

- A person commits capital murder if he or she intentionally commits the murder in the course of committing or attempting

to commit kidnapping, burglary, robbery, aggravated sexual assault, or arson.

- A person commits voluntary manslaughter if he or she committed murder under the immediate influence of sudden passion arising from an adequate cause.

- A person commits involuntary manslaughter if he or she recklessly causes the death of an individual.

4. A drunk woman drove out to the airplane hangar. As she started the engine of her airplane, she did not notice the man standing next to the propeller. The man died as the result of massive head and chest wounds. According to the definitions above, the woman should be charged with:

- A. murder
- B. capital murder
- C. voluntary manslaughter
- D. involuntary manslaughter
- E. none of the above

5. In a desperate attempt to get the bank employees to comply with his demands for money, the suspect shot the bank president in the head. Despite the fact that the suspect received no money, he fled immediately. The bank president died as a result of the gunshot wound. According to the definitions above, the suspect should be charged with:

- A. murder
- B. capital murder
- C. voluntary manslaughter
- D. none of the above

6. A woman discovered that her husband was having an affair. Upon hearing the news, she became violently angry. She spent one week planning the poisoning of her husband's mistress. Upon drinking tea poisoned by the wife, the mistress died. According to the definitions above, the woman should be charged with:

- A. murder
- B. capital murder
- C. *voluntary manslaughter*
- D. none of the above

* Questions 7-12 are based on the following definitions. Do not assume anything to be true that is not stated in the definitions. Select the charge that fits the situation best.

- A person who conducts himself or herself in a way that places another in reasonable concern of receiving bodily harm commits assault.

- A person commits aggravated assault if he or she uses a deadly weapon while committing an assault.

- A person is guilty of battery if he or she purposely, without legal justification, causes bodily harm to an individual.

- A person commits aggravated battery if he or she uses a deadly weapon while committing a battery.

7. John tried to punch Teresa in the face. Teresa ducked just in time to miss John's fist. According to the definitions, John could be charged with:

- A. assault
- B. aggravated assault
- C. battery
- D. aggravated battery
- E. none of the above

8. Michelle told Bob that if he spilled his beer one more time, she would shoot him in the head. Although Michelle waved the gun in Bob's face, she did not shoot him. According to the definitions Michelle has committed:

- A. assault
- B. aggravated assault
- C. battery
- D. aggravated battery
- E. none of the above

9. A man walked up to a woman, placed both hands on her chest, and shoved her to the ground. He has committed:

- A. assault
- B. aggravated assault
- C. battery
- D. aggravated battery
- E. none of the above

10. According to the definitions, the occupation of an individual threatened or injured in an assault and/or battery case has no effect on the charges brought against the suspect.

- A. True
- B. False

11. After a verbal argument in a nightclub, two men walked outside to "discuss" their problem. One man is unarmed and says nothing. The other pulled out a knife with a blade ten inches long. He swung the knife across the other man's face. The injured party bled profusely, and his wound required 30 stitches. According to the definitions, the man with the knife committed:

- A. assault
- B. aggravated assault
- C. battery
- D. aggravated battery
- E. none of the above

12. A young student approached her teacher, pulled out a small gun from her purse, and threatened, "You're going to be sorry you ever crossed my path." According to the definitions, the girl committed:

- A. assault
- B. aggravated assault
- C. battery
- D. aggravated battery
- E none of the above

* For questions 13-24, select the best analogy from the choices listed.

13. MISDEMEANOR is to CRIME as EAGLE is to:

- A. man
- B. hawk
- C. bird
- D. worm

14. CATCH is to APPREHEND as SEARCH is to:

- A. find
- B. seek
- C. discover
- D. overlook

15. LAWLESSNESS is to CHAOS as RULES are to:

- A. contention
- B. dismissal
- C. order
- D. anarchy

16. MOTIVE is to REASON as ALIBI is to:

- A. question
- B. confession
- C. excuse
- D. deceit

17. DISREGARD is to IGNORE as ENFORCE is to:

- A. uphold
- B. avoid
- C. postpone
- D. neglect

18. ILLICIT is to LEGAL as UNLAWFUL is to:

- A. illegal
- B. legitimate
- C. arbitrary
- D. conventional

19. IDENTIFY is to PERCEIVE as RECOGNIZE is to:

- A. liberate
- B. divulge
- C. adhere
- D. discern

20. SUBMIT is to RESIST as SURRENDER is to:

- A. relinquish
- B. withstand
- C. coerce
- D. succumb

21. FLAG is to FREEDOM as CAGE is to:

- A. autonomy
- B. reinforcement
- C. bondage
- D. contain

22. COMPLIANCE is to OBEDIENCE as HONESTY is to:

- A. acquiescence
- B. mendacity
- C. dishonesty
- D. frankness

23. INFRINGE is to INVADE as SURROUND is to:

- A. encircle
- B. connect
- C. block
- D. trespass

24. RESPECT is to HONOR as PERSECUTE is to:

- A. praise
- B. abuse
- C. venerate
- D. delude

25. The floor of the room in which the crime took place was square in shape and 121 square feet in area. How many feet long was each wall in the room?

- A. 2
- B. 11
- C. 21
- D. 60.5

26. Officer Jones kept track of the miles he put on his vehicle. He thought he had driven about 676 miles, but he actually drove 89 miles less. How many miles did he actually drive?

- A. 573
- B. 585
- C. 587
- D. 597

27. Inspector Kincaid confiscated 46 cans of coffee because she suspected they contained cocaine. Upon inspection, each can contained 7 grams of cocaine. How many grams of cocaine did she find?

- A. 282
- B. 293
- C. 302
- D. 322
- E. 342

28. The three bank robbers were 36, 78, and 104 years of age. What was the average age of the bank robbers?

- A. 70
- B. 72.67
- C. 76.67
- D. 78
- E. 78.73

29. The elevator at headquarters can carry 300 pounds of weight. If the computer monitors weigh 60 pounds each, how many computer monitors can the elevator carry at one time?

- A. 5
- B. 15
- C. 30
- D. 50
- E. 55

30. The criminal was 6 feet, 3 inches tall. How many inches tall was he?

- A. 39
- B. 63
- C. 72
- D. 75
- E. 78

31. The four boxes weighed 10, 37, 59, and 74 pounds. What was the total weight of the boxes?

- A. 170 pounds
- B. 171 pounds
- C. 180 pounds
- D. 190 pounds
- E. 192 pounds

32. A jewelry store has been burglarized, and 21 watches with a total value of \$1050.00 have been stolen. What is the average value of the watches?

- A. \$10.50
- B. \$50.00
- C. \$55.00
- D. \$105.50
- E. \$2,050.00

33. While searching a suspected thief, the police officer found a great deal of cash in his pockets. The cash included the following: 30 100-dollar bills, 10 50-dollar bills, and 100 20-dollar bills. How much cash was the suspect carrying ?

- A. \$5,000.00
- B. \$5,500.00
- C. \$6,000.00
- D. \$6,400.00
- E. \$6,800.00

34. An old woman broke into a truck. If she stole 25 portable compact disc players having a total retail value of \$9,998.75, what was the average retail value of each compact disc player?

- A. \$259.95
- B. \$299.95
- C. \$369.95
- D. \$399.95
- E. \$400.00

35. Sergeant Shaw was retiring after 25 years of service. If each of his 53 coworkers contributed \$10, what was the total amount that could be spent on a retirement gift?

- A. \$440.00
- B. \$500.00
- C. \$530.00
- D. \$1,330.00
- E. \$5,300.00

36. State Trooper Morales patrols a 50-mile stretch, one way, of the interstate highway system. In one day he made 2.5 round trips on his patrol stretch. How many miles did he travel?

- A. 125
- B. 205
- C. 250
- D. 2050
- E. 2550

37. Miss Richter was arrested for shoplifting. In order to be released from police custody, she had to pay 10% of the \$5000.00 bail set by the judge. How much did she have to pay?

- A. \$5.00
- B. \$50.00
- C. \$250.00
- D. \$500.00
- E. \$550.00

38. An automobile traveling on an icy road slid 132 inches into a ditch. In preparing the diagram for her accident report, Sergeant Wren needed to state in feet how far the out-of-control vehicle traveled. How many feet did the car slide?

- A. 11
- B. 12
- C. 13.2
- D. 96
- E. 112

39. A homeowner placed the value of items stolen in a burglary at \$1093.00. If he valued the television set at \$405.00, what is the total value of the remaining items?

- A. \$1498.00
- B. \$688.00
- C. \$628.00
- D. \$505.00
- E. \$403.00

40. Officer Williams put 414 miles on his squad car in one five-day week. He put 114 of those miles on his car in one day. On the average, how many miles did he put on the car on each of the other days?

- A. 103.6
- B. 100.0
- C. 82.8
- D. 75.0
- E. 66.0

41. A station wagon was traveling 35 miles per hour in a 20 mph school zone. The driver was fined \$5.00 for each mile per hour over the speed limit. How much was her fine?

- A. \$25.00
- B. \$50.00
- C. \$75.00
- D. \$100.00
- E. \$125.00

42. An arrested man's bail was set at \$3500.00. His wife had to pay 10% of that amount in cash in order for her husband to be released. She only had \$205.00 in her purse. How much more money did she need?

- A. \$45.00
- B. \$135.00
- C. \$145.00
- D. \$153.00
- E. \$155.00

43. Lieutenant Cooper was practicing his aim at the firing range. In one afternoon he shot 600 bullets. He hit the target 85% of the time. How many bullets missed the target?

- A. 60
- B. 75
- C. 85
- D. 90
- E. 95

44. "Thousands of automobiles are stolen each year in this country. Thieves take cars apart, reassemble them, and sell them to dealers. The altered products are sold by dealers who also sell legitimate used cars." The statements indicate that many stolen automobiles are:

- A. recovered by the authorities
- B. taken apart and sold for spare parts
- C. disposed of at auctions
- D. difficult to identify when sold

45. "Intelligence is not the only requirement of a successful supervisor. He or she must be honest and industrious."

According to this statement, a successful supervisor must be:

- A. able to handle problems
- B. as intelligent as he or she is industrious
- C. honest and industrious as well as intelligent
- D. loyal no matter what circumstances exist

* Questions 46-49 are based on the following selection:

source: Campus Law Enforcement Journal; Vol. 22, No. 3, M/J 92
author: Daniel J. Benny

Those in the campus law enforcement and security profession may encounter Satanic activity at some point in the course of an investigation. It is important to recognize signs related to Satanism, as they may have an impact on the possible subject, motive, or identified physical evidence related to the crime.

Satanism is the worship of the Devil; the choice of evil over good. The practice of Satanism in itself is not illegal; in fact, it is protected by the United States Constitution under freedom of religion. However, certain activities related to that worship, including the use of drugs, trespassing, criminal mischief, arson, animal mutilation, kidnapping, rape, child abuse, and murder are illegal. The important fact to remember in investigating criminal acts related to Satanism is to focus on the crime, rather than on the practice of Satanism itself.

Satanism may provide the motive, but as an investigator, you must be impartial and focus on the facts.

In order to identify Satanic involvement in criminal activity, the investigator must have knowledge of and be able to identify Satanic paraphernalia which may be found at the crime scene. Indicators of Satanic involvement include the use of Satanic symbols, such as the upside-down pentagram or Baphomet, which is a five-pointed star representing the head of a goat, the inverted cross, the numbers 666 taken from the Book of Revelation as the sign of the Beast or Satan, vandalized Christian artifacts or the use of candles or candle drippings, and animal mutilations.

The existence of hard core Satanic cults on college campuses is rare. However, it is important that those in the campus law enforcement and security profession be aware of the possibility of serious criminal activity on the part of those involved in Satanism. Those individuals may attempt to recruit college students to join their groups or lure them into participation in activities as victims of Satanic rituals. Through education and gaining an understanding of this complex phenomenon, the campus security professional will be able to perform a professional and impartial investigation of Satanic-related crimes on campus.

46. Why is Satanism legal?

- A. freedom of speech
- B. freedom of expression
- C. freedom to bear arms
- D. freedom of religion

47. Which of the following is not a typical illegal practice of Satanists?

- A. arson
- B. prostitution
- C. kidnapping
- D. murder

48. What should the focus of police investigations be?

- A. the immoral beliefs of the Satanists
- B. the criminal acts of the Satanists
- C. the criminal intentions of the Satanists
- D. all of the above

49. Which of the following is not a typical Satanic symbol?

- A. a burned cross
- B. a Baphomet
- C. a black cat
- D. an inverted cross

* Questions 50 and 51 are based on the following selection:

source: Campus Law Enforcement Journal; Vol. 22, No. 4, J/A
1992 author: Ann P. Glavin

The Massachusetts Institute of Technology (MIT) Campus Police experience with bicycle patrols has thus far indicated that there are virtually no negative aspects to this increasingly popular patrol method. Although one might expect injuries and traffic accidents to be a primary concern, especially on urban campuses, this was not the case for MIT. During the six months of bicycle patrol operation, only one minor injury was reported. In addition, there were no traffic accidents.

The benefits of police on bikes are many: quicker access to areas difficult to reach by cruiser, health benefits for bicycle patrol officers, positive morale benefits for participating officers by breaking the monotony of foot and cruiser patrol, low equipment maintenance costs, and an environmentally sound patrol method.

Most important in the MIT experience is the increased person-to-person contact with members of the MIT community and positive community relations benefits. On a college campus the need for visibility of the campus police in deterring crime is both a community desire and a police

management objective. Bicycle patrols make it an easy and successful achievement.

50. Which of the following did the MIT bicycle patrols experience?

- A. high cost
- B. monotony
- C. minor injury
- D. lost time from duty

51. Which of the following is an implied benefit of bicycle patrols?

- A. more free time
- B. enhanced community relations
- C. increased visibility
- D. increased fitness of officers

* Questions 52-55 are based on the following selection:

source: FBI Enforcement Bulletin, June 1988 author: Lee P. Brown

A "crack" or freebase cocaine house can be defined as any building where crack and other illegal narcotics are sold, which has been fortified to delay entry by police officers armed with search warrants. The extent of fortification varies. Some houses have steel doors, steel bars on windows, and four-foot iron posts embedded around the perimeter of the property to prevent vehicles from ramming the building. Others have the appearance of private clubs which are actually fronts for drug sales and other illegal activities. The purpose of delayed entry is to give the occupants of the house time to dispose of or conceal the drugs.

When a crack house is located in a neighborhood, the problem is not a problem only for the police but for the area residents who experience a deterioration in their quality of life. Residents know that drugs are being sold from the crack houses. The inability of the police to shut the houses down

permanently has made the residents feel a sense of lawlessness and crime in their neighborhood. This perception is a very serious aspect of the crack house problem.

In addition to more traditional enforcement strategies such as the use of informants and undercover police officers, some task forces utilize some nontraditional strategies. Some of these strategies are the following: place a saturation of uniformed patrols in the immediate vicinity of the crack house, park marked patrol cars in front of the crack house, question people in the vicinity of the crack house who are reasonably suspected to be engaged in criminal activity, and notify the property owner that continued violations may result in appropriate action being taken against the property.

Nontraditional strategies have proven to be most effective against the operators of single crack houses. However, the traditional enforcement strategies are still useful in the fight against the operators of multiple crack houses. Because each crack house is different and is run differently, the best solution is to cater the strategies to the situation. Only through concentrated, strategic, and intelligent enforcement activities can the problem of crack houses be eradicated.

52. Which of the following is a cue regarding the existence of a crack house?

- A. high chain-link fence
- B. dark curtains on the windows
- C. no sidewalk
- D. no lights on

53. According to the passage, why are crack houses fortified?

- A. to delay entry of police and other unwelcomed visitors
- B. to make the type of business obvious to potential customers
- C. to provide protection from gangs
- D. both a and c

54. The passage implies that crack houses blend in with neighboring homes, which makes them difficult to detect in residential areas.

- A. True
- B. False

55. Which of the following is not a nontraditional strategy for combating crack houses?

- A. intimidating potential customers
- B. placing more undercover police officers in the area
- C. placing more police cars in the vicinity
- D. increasing the number of uniformed police officers on area patrol

* Questions 56-58 are based on the following selection:

source: FBI Law Enforcement Bulletin, May 1992 authors: Michael H. West, D.D.S. and Robert E. Barsley, D.D.S., J.D.

The same ultraviolet (UV) rays that cause people to get sunburns in the summer also help to catch and prosecute criminals. Researchers are discovering that these rays can literally "cast a new light" on evidence that might not even be detected using conventional investigative techniques. While (UV) technology is still in its early stages, it has already helped to solve crimes and is proving to be a significant development in the field of forensic research.

How can UV or "invisible" light be used in law enforcement? One application is in the analysis of bite marks on human skin. UV light provides more detail and contrast to an injured area than standard lighting techniques. This discovery led to the development of innovative techniques for UV photography.

Thus far, the photos produced by UV imaging have proven useful. Such photographs show wounds in greater detail than would be possible with conventional photographic equipment. UV photographs reveal images of wounds that could not be seen by the naked eye. Therefore, UV light allows investigators and

forensic researchers to examine clues and recover evidence that could not have been detected previously. While the application of UV light is still a relatively new field, it promises to be an indispensable tool for law enforcement.

56. Which of the following is a probable application of UV technology?

- A. examining the skin of a burn victim
- B. examining the wound of someone involved in a knife fight
- C. examining a bullet wound
- D. examining the bite on a rape victim

57. Which of the following best explains the importance of UV technology in law enforcement?

- A. It is another tool which helps determine the time of death.
- B. Natural fibers at the scene of the crime can be analyzed.
- C. Evidence that was once difficult to detect is now available for analysis.
- D. none of the above

58. If conventional photography is less expensive than UV photography, when is

UV technology most likely to be employed?

- A. in cases involving multiple suspects
- B. in all rape cases
- C. in cases lacking substantial physical evidence
- D. in strangulation investigations

* Questions 59-63 are based on the following selection:

source: Texas Police Journal, Vol. 40, No. 9, October 1992

author: Lynn P. Rehm, Ph.D.

Depression can be a problem in any stressful work setting. Law enforcement can be among the most stressful of settings. The patterns of stress most commonly associated with depression involve chronic stresses over a period of time, and situations where punishing experiences seem frequent and

rewards few. The daily encounters of the police officer often seem to fit these patterns. Recognizing depression, however, is not always a simple task.

Many people can list the major mental changes associated with depression such as sadness, hopelessness, helplessness, low self-esteem, guilt, difficulty concentrating, and thoughts of death or suicide. Depression also has recognizable physical symptoms such as fatigue, sleep difficulties, appetite and weight loss (or sometimes overeating and weight gain), and either nervous restlessness or apathy.

Unfortunately, knowing the symptoms is not all that is involved in detecting depression. People who are depressed do not always show all of the known symptoms. Many show subtler symptoms that are more difficult to notice. Sometimes a person's withdrawal from friends and coworkers can be a powerful clue. What is important to detect is change. Other behaviors like unwarranted pessimism, intense self-criticism, and periods of uncontrolled weeping are clues, too.

Depression occurs in all degrees of severity. Clinical psychologists consider depression to be a disorder when it interferes with the person's life. At this level, people often find it difficult to solve problems on their own and may require professional help. Severely depressed individuals find themselves unable to function and incapable of dealing with everyday duties. Intervention is a necessity at this level.

The good news is that depression is treatable. Advances in drug treatments and effective psychological approaches are helpful not only in overcoming depression but also in reducing the likelihood of recurrence. Professionals can treat depression in short periods of time, and treatment is affordable. By recognizing symptoms and signs of depression, people can encourage depressed individuals to seek available help.

59. Which of the following is not a common symptom of depression?

- A. low self-esteem
- B. concentration deficits
- C. unusual concern for others
- D. fatigue

60. Professionals consider depression to be a clinical disorder when:

- A. it results in a negative attitude toward life
- B. it makes people feel helpless
- C. it affects everyday life in a maladaptive way
- D. it causes someone to lose sleep

61. Which types of change are always seen in depressed people?

- A. mental and relationship
- B. interest and physical
- C. both a and b
- D. neither a nor b

62. Which of the following summarizes the main points of the selection?

- A. recognition of symptoms, mental and physical changes, and affordability of treatment
- B. symptoms, levels of severity, and treatment approaches
- C. causes, mental and physical changes, and benefits of treatment
- D. recognition of symptoms, severity differences, and benefits of treatment

63. The passage implies that police officers are more likely to become depressed than those individuals whose occupations involve less stress.

- A. True
- B. False

* The following sentences are missing one word. Select the correct spelling of the word that completes the sentence.

64. The suspect sat in the _____ room.

- A. interrogation
- B. interrogation
- C. enterogation
- D. enterrogation

65. There were red stains on the suspect's _____.

- A. collar
- B. coller
- C. colar
- D. coler

66. The store manager called the police because he saw a _____ prowler.

- A. suspishous
- B. suspiscious
- C. suspicious
- D. suspischus

67. The wounded gang member was _____ across the parking lot.

- A. stagering
- B. staggering
- C. staggerring
- D. stagerring

68. The bartender refused to serve more beer to the _____ patron.

- A. beligerent
- B. beligerant
- C. belligerent
- D. belligerrent

69. Upon arrival at the hospital, the car accident victim was _____.

- A. comatouse
- B. comatose
- C. cometose
- D. comeatose

70. The gas station _____ was asleep in the back room while the theft took place.

- A. atendent
- B. attendent
- C. attendant
- D. atendant

71. The child who fell from the tree _____ several bones in his arms and legs.

- A. frachtured
- B. fractured
- C. fractoured
- D. fracstured

72. The man argued that he had his girlfriend's _____ to withdraw the contents of her safe deposit box.

- A. consent
- B. cunsent
- C. conssent
- D. conscent

73. Mr. Richardson received his _____ early this morning.

- A. supena
- B. subpena
- C. subpoena
- D. supoena

74. The police _____ the revolver during the struggle.

- A. ceized
- B. ceased
- C. seized
- D. seazed

75. It was amazing that the baby was still _____ after being hurled from the vehicle.

- A. conscious
- B. conshous
- C. conscience
- D. conschious

76. A _____ took place at three o'clock in the morning.

- A. burgary
- B. burglary
- C. burglary
- D. burglarize

77. Officer Morris apprehended the _____ in the alley.

- A. crimanel
- B. criminal
- C. crimanal
- D. crimenel

78. The drug dealer destroyed the _____ before the police arrived.

- A. evedance
- B. evidence
- C. evidance
- D. evedence

79. The _____ phone caller terrorized Mrs. Jorgensen.

- A. obsean
- B. obscene
- C. obsene
- D. obscean

80. _____ is a depressant drug.

- A. Alcolhol
- B. Alchole
- C. Alcohol
- D. Alcahol

81. Mrs. Scanlon's motive for shooting her husband is not _____.

- A. apparent
- B. aperant
- C. apperant
- D. apperent

82. They conducted the homicide _____ for three weeks.

- A. investigation
- B. investagation
- C. investegation
- D. invesigation

83. Officer Perkins _____ the information yesterday.

- A. received
- B. recieved
- C. rescieved
- D. reiceved

84. The assault _____ last night.

- A. occured
- B. accurred
- C. occurred
- D. ocurred

85. It is _____ that the police will encounter protesters at the rally.

- A. posible
- B. possible
- C. posable
- D. possable

86. The student was arrested for disturbing the _____.

- A. peace
- B. piece
- C. peice
- D. pease

87. Johnny Mottola's _____ is "The Jackhammer."

- A. alles
- B. alies
- C. alias
- D. ailiass

88. "Minor violations of the traffic rules will not be tolerated." As used in this sentence, the word "tolerated" means most nearly:

- A. investigated
- B. reported
- C. punished
- D. permitted

89. "The chief will defer action on that matter." The word "defer" most nearly means:

- A. cancel
- B. consider
- C postpone
- D. expedite

90. Which of the following sentences contains no error in grammar, spelling, punctuation, or word usage?

- A. This matter is between you and me.
- B. Place each item in it's original position.
- C. The truck remained stationery all day long.
- D. We must decide whose responsible.

* For each of the following items, the first word is in capital letters. Select the letter corresponding to the word that comes closest to the meaning of the word in capital letters.

91. SPACE

- A. school
- B. noon
- C. captain
- D. room
- E. board

92. BROADEN

- A. efface
- B. make level
- C. elapse
- D. embroider
- E. widen

93. EMANATE

- A. populate
- B. free
- C. prominent
- D. rival
- E. come

94. EDIBLE

- A. auspicious
- B. eligible
- C. fit to eat
- D. sagacious
- E. able to speak

95. ANIMOSITY

- A. hatred
- B. animation
- C. disobedience
- D. diversity
- E. friendship

96. PACT

- A. puissance
- B. remonstrance
- C. agreement
- D. skillet
- E. pressure

97. CLOISTERED

- A. miniature
- B. bunched
- C. arched
- D. malady
- E. secluded

98. CAPRICE

- A. value
- B. a star
- C. grimace
- D. whim
- E. inducement

99. ACCUSTOM

- A. disappoint
- B. customary
- C. encounter
- D. get used to
- E. business

100. ALLUSION

- A. reference
- B. dream
- C. eulogy
- D. illusion
- E. aria

Appendix CRefugio Test

Instructions: There is only one correct answer for each question. When you select a response, mark the corresponding box on the answer sheet. Be sure that your shading is dark and neat.

1. Which of the following statements correctly states Trooper Lopez's reason for stopping the white vehicle?
 - a. Trooper Lopez noticed that the vehicle was swerving.
 - b. Trooper Lopez detected a faulty head lamp on the vehicle.
 - c. Trooper Lopez thought the vehicle appeared to be speeding.
 - d. Trooper Lopez thought the vehicle was stolen.
2. Select the correct spelling of that to which the interviewer was referring when he mentioned "buyer's tags:"
 - a. a temporary lisence
 - b. a temperary license
 - c. a temporary license
 - d. a temperary liscense
3. Which of the following did not make Trooper Lopez suspicious?
 - a. the blood on the front seat
 - b. the towel on the passenger's shoulder
 - c. the glances exchanged between the driver and the passenger
 - d. the nervous behavior of the driver
4. Select the correct spelling of the odor Trooper Lopez detected:
 - a. marahuana
 - b. marijuana
 - c. marehuana
 - d. marajuana

5. Once the occupants were out of the vehicle, which of the following lists best describes Trooper Lopez's priorities?
- stay in control, stay safe, and get help
 - stay in control, hide beneath the car, and wait for help
 - draw the weapon, take aim, and get cover immediately
 - draw the weapon, stay safe, and call for help
6. Which of the following correctly spells the reason for which Trooper Lopez wanted to avoid the headlight beams?
- their ilumination
 - there ilumination
 - there ilumenation
 - their illumination
7. Which of the following sentences contains no error in grammar, spelling, punctuation, or word usage?
- Trooper Lopez was in high physical condition with a great attitude and has maximum alertness.
 - Trooper Lopez's high physical condition enables him to maintain a positive attitude and exceptional alertness.
 - Trooper Lopez, with high physical condition, kept alert and his attitude.
 - Due to his high physical condition, Trooper Lopez kept his positive mental attitude with his alertness.
8. How long did it take Trooper Lopez to shove the assailant, yell, clear the weapon, and return fire on the run?
- 1.33 seconds
 - 2.66 seconds
 - 3.66 seconds
 - 6.22 seconds
9. The bullet grazed which part of Trooper Lopez's body?
- his abdomen
 - his hip
 - his thigh
 - his ear

10. What was Trooper Lopez's hit percentage if he made eight hits after firing 14 rounds?

- a. 48%
- b. 52%
- c. 57%
- d. 62%

11. Which of the following is not a point made by Trooper Lopez?

- a. wear body armor
- b. stay in good physical condition
- c. keep your weapon armed
- d. stay mentally prepared

12. If Trooper Lopez were traveling due west and decided to make a 45-degree turn to the left, in which direction would he then be traveling?

- a. southeast
- b. southwest
- c. northeast
- d. northwest

Appendix D

Instructions for the Writing Sample

After viewing the video, subjects were asked to follow these instructions: Write a few paragraphs giving as many important details as you can about the situation you have just witnessed. Use grammatically correct, complete sentences to convey your points. Write or print legibly only on this side of the paper.

Table 1

The University of Texas System Police Departments That
Participated in the Development Phase of the Entry-Level
Examination

1. The University of Texas at Austin
2. The University of Texas at Arlington
3. The University of Texas at Dallas
4. The University of Texas at El Paso
5. The University of Texas at San Antonio
6. The University of Texas at Tyler
7. The University of Texas Health Center at Tyler
8. The University of Texas Science Center at San Antonio
9. The University of Texas - Houston Institutions
10. The University of Texas Medical Branch at Galveston
11. The University of Texas - Pan American
12. The University of Texas of the Permian Basin
13. The University of Texas Southwestern Medical Center at Dallas
14. The University of Texas System Police Headquarters at Austin

Table 2

The Linkage Between the Eight Cognitive Abilities and the
Subtest Areas of the Entry-Level Examination

Cognitive Ability	Related Subtests
Deductive Reasoning	1 and 3
Inductive Reasoning	2 and 4
Information Ordering	6 and 7
Memorization	6
Problem Sensitivity	3 and 8
Spatial Orientation	8
Verbal Comprehension	2 and 5
Verbal Expression	5 and 9

Note. This table indicates the subtests which primarily evaluate the cognitive abilities.

Table 3

The University of Texas System Police Departments that
Provided Incumbents for Testing

<u>Institution</u>	<u>Number of Officers Tested</u>
Austin	29
Arlington	21
Dallas	6
El Paso	11
San Antonio	10
Health Science Center at San Antonio	6
Houston Institutions	43
Medical Branch at Galveston	14
Pan American	6
Southwestern Medical Center at Dallas	13
Police Academy at Austin	38
<u>Total Number of Officers Tested</u>	<u>197</u>

Table 4

Race and Gender Compositions of the Cadet Sample, Incumbent Sample, and the UTSP Force

<u>Race and Gender</u>	<u>Cadet Sample</u>	<u>Incumbent Sample</u>	<u>UTSP Force</u>
Black Females	1	7	8
Black Males	5	22	36
Hispanic Females	1	4	8
Hispanic Males	6	19	31
White Females	6	18	28
White Males	19	89	160
Total Females	8	29	43
Total Males	30	130	230
TOTAL	38	159	273

Table 5

Means, Variances, and Correlations Among the Dependent
Variables and the Independent Variables

Independent Variable Mean and (Variance)	Correlations	Dependent Variable Mean and (Variance)
82.95 (83.25) 79.71 (112.09)	Digit x Academic Avg	
	Cadet Sample: 0.48	85.21 (28.95)
	Incumbent Sample: 0.18	84.73 (34.76)
82.95 (83.25) 79.71 (112.09)	Digit x Final Avg	
	Cadet Sample: 0.42	84.95 (28.04)
	Incumbent Sample: 0.16	84.87 (29.14)
82.10 (88.09) 81.03 (107.51)	Main II Test x Academic Avg	
	Cadet Sample: 0.77	85.21 (28.95)
	Incumbent Sample: 0.59	84.73 (34.76)
82.10 (88.09) 81.03 (107.51)	Main II Test x Final Avg	
	Cadet Sample: 0.73	84.95 (28.04)
	Incumbent Sample: 0.57	84.87 (29.14)
75.00 (63.30) 74.30 (82.62)	Main II - Vocab x Academic Avg	
	Cadet Sample: 0.77	85.21 (28.95)
	Incumbent Sample: 0.57	84.73 (34.76)
75.00 (63.30) 74.30 (82.62)	Main II - Vocab x Final Avg	
	Cadet Sample: 0.73	84.95 (28.04)
	Incumbent Sample: 0.54	84.87 (29.14)
75.94 (45.54) 70.67 (110.13)	Reading x Academic Avg	
	Cadet Sample: 0.04	85.21 (28.95)
	Incumbent Sample: -.13	84.73 (34.76)
75.94 (45.54) 70.67 (110.13)	Reading x Final Avg	
	Cadet Sample: -.03	84.95 (28.04)
	Incumbent Sample: -.15	84.87 (29.14)
10.11 (1.88) 9.71 (3.75)	Refugio Test x Academic Avg	
	Cadet Sample: 0.34	85.21 (28.95)
	Incumbent Sample: 0.46	84.73 (34.76)
10.11 (1.88) 9.71 (3.75)	Refugio Test x Final Avg	
	Cadet Sample: 0.34	84.95 (28.04)
	Incumbent Sample: 0.43	84.87 (29.14)

(table continues)

Independent Variable Mean and (Variance)	Correlations	Dependent Variable Mean and (Variance)
11.61 (15.16)	Spatial x Academic Avg	85.21 (28.95)
11.01 (14.72)	Cadet Sample: 0.49	84.73 (34.76)
	Incumbent Sample: 0.29	
11.61 (15.16)	Spatial x Final Avg	84.95 (28.04)
11.01 (14.72)	Cadet Sample: 0.40	84.87 (29.14)
	Incumbent Sample: 0.22	
6.76 (2.29)	Vocabulary x Academic Avg	85.21 (28.95)
6.86 (3.10)	Cadet Sample: 0.53	84.73 (34.76)
	Incumbent Sample: 0.55	
6.76 (2.29)	Vocabulary x Final Avg	84.95 (28.04)
6.86 (3.10)	Cadet Sample: 0.40	84.87 (29.14)
	Incumbent Sample: 0.53	
8.28 (2.53)	Writing x Academic Avg	85.21 (28.95)
8.19 (3.03)	Cadet Sample: 0.34	84.73 (34.76)
	Incumbent Sample: 0.15	
8.28 (2.53)	Writing x Final Avg	84.95 (28.04)
8.19 (3.03)	Cadet Sample: 0.41	84.87 (29.14)
	Incumbent Sample: 0.14	

Table 6

Test Statistics (Main II Test and Refugio Test)

Examinees	<i>n</i>	Mean	Standard Deviation	Reliability
(Main II Test)				
<u>Cadets</u>	21	81.57	9.39	0.89
<u>Incumbents</u>				
All	159	81.03	10.78	0.90
Black	29	75.76	9.56	0.87
Female	29	82.03	7.43	0.81
Hispanic	23	74.41	13.15	0.93
Male	130	80.71	10.90	0.91
White	107	84.02	8.81	0.87
(Refugio Test)				
<u>Cadets</u>	38	10.11	1.37	0.38
<u>Incumbents</u>				
All	159	9.71	1.94	0.59
Black	29	9.35	1.88	0.55
Female	29	9.72	1.67	0.44
Hispanic	23	8.97	2.22	0.65
Male	130	9.72	2.01	0.62
White	107	10.06	1.79	0.56

Note. The Cronbach alpha formula was used to calculate the reliability measure.

Table 7
Correlation Matrix (Incumbent Sample)

Items	Sample					
	All	Black	Female	Hispanic	Male	White
Digit x Acad Avg	0.18	0.01	-.06	0.27	0.22	0.13
Digit x Final Avg	0.16	0.09	-.02	0.18	0.20	0.08
Main II x Acad Avg	0.59	0.20	0.33	0.57	0.64	0.60
Main II x Final Avg	0.57	0.24	0.23	0.52	0.62	0.53
Main II - Vocab x Acad Avg	0.57	0.19	0.32	0.60	0.62	0.56
Main II - Vocab x Final Avg	0.54	0.23	0.24	0.54	0.59	0.49
Reading x Acad Avg	-.13	-.05	-.11	-.18	-.13	-.14
Reading x Final Avg	-.15	-.08	-.04	-.26	-.16	-.12
Refugio x Acad Avg	0.46	0.35	0.20	0.47	0.51	0.43
Refugio x Final Avg	0.43	0.36	0.16	0.43	0.47	0.36
Spatial x Acad Avg	0.29	0.20	0.14	0.21	0.33	0.12
Spatial x Final Avg	0.22	0.16	-.07	0.26	0.27	0.03
Vocab x Acad Avg	0.55	0.14	0.17	0.47	0.61	0.56
Vocab x Final Avg	0.53	0.14	0.08	0.45	0.59	0.52
Writing x Acad Avg	0.15	0.17	0.26	0.19	0.14	0.12
Writing x Final Avg	0.14	0.17	0.21	0.19	0.13	0.12

Note. Abbreviations Used:

Acad Avg = Academic Average

Digit = Digit Substitution Task

Final Avg = Final Average

Main II = Main II Test

Main II - Vocab = Main II Test (without vocabulary subsection)

Reading = Reading Ease (of writing sample)

Refugio = Refugio Test

Spatial = Spatial Reasoning Task

Writing = Writing (Grade) Level (of writing sample)

Table 8

Test Score Differences by Race, Sex, and Group

Group	<i>n</i>	Mean	Standard Deviation	<i>t</i> - test	Significance Level (.01)
White	107	84.02	8.81	--	--
Black	29	75.76	9.56	+2.34	not significant
Hispanic	23	74.41	13.15	+1.93	not significant
Male	130	80.71	10.90	--	--
Female	29	82.03	7.43	-0.29	not significant
Incumbents	159	81.03	10.78	--	--
Cadets	21	81.57	9.39	0.00	not significant

Table 9

Attrition Summary for UTSP Academy Classes 46 Through 55

Class	Number of Officers	Class Average	Number of Resignees	Resignee Average	Delta
46	18	85.91	7	88.33	+2.39
47	24	85.40	12	85.48	+0.08
48	22	86.61	6	89.77	+3.16
49	17	88.57	6	88.38	-0.19
50	14	86.00	4	88.73	+2.73
51	23	78.97	2	81.50	+2.53
52	24	82.91	2	81.79	-1.12
53	23	83.72	2	87.11	+3.39
54	15	83.32	N/A	N/A	N/A
55	18	82.90	N/A	N/A	N/A

Table 10

Race and Gender of the Resignees from UTSP Academy Classes
46 Through 53 (Compared to UTSP Composition Percentages)

<u>Race and Gender of Cadets</u>	<u>Number of Resignees</u>	<u>Percent of Cadets who Resigned</u>	<u>Percent of Total Resignees</u>	<u>Percent of Total UTSP Force</u>
<u>Black</u>	8	4.06	19.05	34.12
Females	2	1.02	4.76	2.93
Males	6	3.05	14.29	13.19
<u>Hispanic</u>	8	4.06	19.05	14.29
Females	0	0	0	2.93
Males	8	4.06	19.05	11.36
<u>White</u>	26	13.20	61.90	68.87
Females	2	1.02	4.76	10.26
Males	24	12.18	57.14	58.61
<u>All (n)</u>	42	21.32	100.00	100.00
Females (32)	4	2.03	9.52	15.75
Males (165)	38	19.29	90.48	84.25
<u>Total (197)</u>	42	21.32	90.48	84.25

VITA

Anne Caroline Parker was born in Burlington, North Carolina, on December 6, 1968, the daughter of Eileen Riley Parker and Louis Francis Parker. After graduating from Rosary High School in 1986, she began her military career by attending the United States Air Force Academy in Colorado Springs, Colorado. In May, 1990, she received a Bachelor of Science degree from the Air Force Academy. On graduation day, she was commissioned as a second lieutenant in the United States Air Force. She then moved to Lubbock, Texas, to attend Undergraduate Pilot Training at Reese Air Force Base. Following flight training, she worked in the Consolidated Base Personnel Office at Reese Air Force Base as the deputy section chief for Personnel Utilization. In January, 1992, she entered The Graduate School of The University of Texas.

Paper presented: "Gender and Attitudes as Predictors of Video Game Performance and Satisfaction," Thirteenth Symposium of Psychology in the Department of Defense, April, 1992, United States Air Force Academy, Colorado.

Permanent address: 280 Lakeview
New Braunfels, Texas 78130

This thesis was typed by the author.

References

- Alwin, D. F. (1988). Cohort differences in vocabulary knowledge in the United States. Ann Arbor, MI: University of Michigan, Department of Sociology and Institute for Social Research.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). Standards for educational and psychological tests. Washington, DC: American Psychological Association.
- Arvey, R. D., & Faley, R. H. (1988). Fairness in selecting employees (2nd ed.). Reading, MA: Addison-Wesley.
- Blum, R. H. (Ed.). (1964). Police selection. Springfield, IL: Charles C. Thomas.
- Brown Maroney. (1991). Law watch. Austin, TX: Brown Maroney & Oaks Hartline.
- Cascio, W. F. (1987). Costing human resources: The financial impact of behavior in organizations (2nd ed.). Boston: PWS-Kent.
- Civil Rights Act of 1991, 42 U. S. C. 2002e-2 (1991).
- Colarelli, N. J., & Siegel, S. M. (1964). A method of police personnel selection. Journal of Criminal Law, Criminology, and Police Science, 55, 287-289.

- Cook, M. (1988). Personnel selection and productivity.
New York: John Wiley & Sons.
- Cronbach, L. J. (1970). Essentials of psychological testing
(3rd ed.). New York: Harper & Row.
- Crosby, A., Rosenfeld, M., & Thornton, R. F. (1979).
Development of a written test. In C. D. Spielberger (Ed.),
Police selection and evaluation: Issues and techniques
(pp. 143-153). Washington: Hemisphere.
- Dobrish, C., Wolff, R., & Zevnik, B. (1984). Hiring the right
person for the right job. New York: Franklin Watts.
- Dunnette, M. D. (1966). Personnel selection and placement.
Belmont, CA: Wadsworth.
- Equal Employment Opportunity Commission, Civil Service
Commission, Department of Labor, & Department of
Justice (1978). Uniform guidelines on employee
selection procedures, Federal Register, 43, 38295-
38309.
- Fabricatore, J. M. (1979). Preentry assessment and training:
Performance evaluation of police officers. In C. D.
Spielberger (Ed.), Police selection and evaluation: Issues and
techniques (pp. 77-85). Washington: Hemisphere.

- Feild, H. S., Holley, W. H., & Buckner, K. E. (1990). Personnel selection in police and fire departments: A study of employment discrimination case characteristics and outcomes. Labor Law Journal, 9, 622-632.
- Feltham, R. (1988). Validity of an assessment centre: A 1-19-year follow-up. Journal of Occupational Psychology, 61, 129-144.
- Fleishman, E. A. (1975). Toward a taxonomy of human performance. American Psychologist, 30, 1127-1149.
- Froemel, E. C. (1979). Objective and subjective measures of performance. In C. D. Spielberger (Ed.), Police selection and evaluation: Issues and techniques (pp. 87-111). Washington: Hemisphere.
- Gatewood, R. D., & Feild, H. S. (1990). Human resource selection. Hinsdale, IL: The Dryden Press.
- Grider, K., & Bailey, T. A. (1992). The civil rights act of 1991: Congress reverses the supreme court. Texas Bar Journal, 1, 37-40.
- Guion, R. M. (1966). Employment tests and discriminatory hiring. Industrial Relations, 5, 20-30.
- Guion, R. M. (1993). The need for change: Six persistent themes. In N. Schmitt, W. C. Borman, & Associates (Eds.), Personnel selection in organizations (pp. 481-515). San Francisco: Jossey-Bass.

- Hirsh, H. R., Northrop, L. C., & Schmidt, F. L. (1986). Validity generalization results for law enforcement occupations. Personnel Psychology, 39, 399-420.
- Hunter, J. E. (1983). A causal analysis of cognitive ability, job performance, and supervisory ratings. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), Performance measurement and theory (pp. 257-266). Hillsdale, NJ: Erlbaum.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. Psychological Bulletin, 96, 72-98.
- Hunter, J. E., & Schmidt, F. L. (1989). Meta-analysis: Facts and theories. In M. Smith & I. T. Robertson (Eds.), Advances in selection and assessment (pp. 203-216). New York: John Wiley & Sons.
- International Personnel Management Association. (1988). Development and validation of the A-2 entry-level police officer examination. Unpublished manuscript.
- International Personnel Management Association. (1992). Development and validation of the A-3 police officer selection test for the international personnel management association (IPMA). Unpublished manuscript.
- Jurkanin, T. J. (1989). Job task analysis. FBI Law Enforcement Bulletin, 11, 10-13.

- Kellogg, C. E., & Morton, N. W. (1962). Revised beta examination. New York: Psychological Corporation.
- Kellogg, C. E., & Morton, N. W. (1962). Revised beta examination manual. New York: Psychological Corporation.
- Krzystofiak, F., Cardy, R., & Newman, J. (1988). Implicit personality and performance appraisal effectiveness. Journal of Applied Psychology, 73, 515-521.
- Landy, F. J. (1989). Psychology of work behavior (4th ed.). Pacific Grove, CA: Brooks/Cole.
- Martin, S. L., & Raju, N. S. (1992). Determining cutoff scores that optimize utility: A recognition of recruiting costs. Journal of Applied Psychology, 77(1), 15-23.
- Morrison & Foerster. (1992). Employment law commentary. San Francisco, CA: Author.
- Moynihan, J. B. (1992). Employment discrimination law in the federal sector: After the civil rights act of 1991. Texas Bar Journal, 5, 450-456.
- Pynes, J. E., & Bernardin, H. J. (1989). Predictive validity of an entry-level police officer assessment center. Journal of Applied Psychology, 74(5), 831-833.
- Roznowski, M. (1987). Use of test manifesting sex differences as measures of intelligence: Implications for measurement bias. Journal of Applied Psychology, 72, 480-483.

- Russell, J. S. (1984). A review of fair employment cases in the field of training. Personnel Psychology, 37, 261-276.
- Sackett, P. R., & Arvey, R. D. (1993). Selection in small n settings. In N. Schmitt, W. C. Borman, & Associates (Eds.), Personnel selection in organizations (pp. 418-447). San Francisco: Jossey-Bass.
- Sauls, J. G. (1992). The civil rights act of 1991: New challenges for employers. FBI Law Enforcement Bulletin, 9, 25-31.
- Schmidt, F. L. (1993). Personnel psychology at the cutting edge. In N. Schmitt, W. C. Borman, & Associates (Eds.), Personnel selection in organizations (pp. 497-515). San Francisco: Jossey-Bass.
- Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). The impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings on job performance. Journal of Applied Psychology, 71, 432-439.
- Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. Personnel Psychology, 37, 407-422.

Schmitt, N., & Landy, F. J. (1993). The concept of validity. In N. Schmitt, W. C. Borman, & Associates (Eds.), Personnel selection in organizations (pp. 275-309). San Francisco: Jossey-Bass.

Schmitt, N., & Robertson, I. (1990). Personnel selection. Annual Review of Psychology. 41, 289-319.

Smart, B. D. (1983). Selection interviewing: A management psychologist's recommended approach. New York: Wiley.

Spielberger, C. D., Ward, J. C., & Spaulding, H. C. (1979). A model for the selection of law enforcement officers. In C. D. Spielberger (Ed.), Police selection and evaluation: Issues and techniques (pp. 11-24). Washington: Hemisphere.

Thorndike, R. L. (1962). Personnel selection: Test and measurement techniques. New York: John Wiley & Sons.

Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen, E. P. (1991). Measurement and evaluation in psychology and education (4th ed.). New York: Macmillan.

Washington v. Davis, 96 S. Ct. 2040 (1976).

Zeidner, J., & Johnson, C. D. (1991). The economic benefits of predicting job performance: Selection utility (Vol. 1). New York: Praeger.