

Association of Firearm and Toolmark Examiners

Occupational Analysis And Certification Examination Development

March 2000

Report Prepared

by

Brian G. Moritsch
Cooperative Personnel Services
191 Lathrop Way, Suite A
Sacramento, CA 95815
(916) 263-3600



Association of Firearm and Toolmark Examiners

Report of Occupational Analysis and Certification Examination Development

Introduction

In January of 1998, the Association of Firearm and Toolmark Examiners (AFTE) entered into a contract with Cooperative Personnel Services (CPS) to develop a series of content-valid certification examinations in three different competency areas: firearm, gunshot residue, and toolmark evidence examination and identification. Each certification examination developed under this contract consisted of a written examination and a practical examination. Two versions of the written and practical certification examinations in each competency area were developed. The purpose of this report is to document the steps taken by CPS and AFTE to develop content-valid certification examinations for the Association of Firearm and Toolmark Examiners. The occupational analysis and test development procedures described in this report were performed to ensure that the resulting AFTE certification examinations are in compliance with applicable professional and legal standards, such as: the Civil Rights Act of 1991; the Americans with Disabilities Act (ADA); Age Discrimination Legislation; the *Uniform Guidelines for Employee Selection Procedures*; the *Principles for the Validation and Use of Personnel Selection Procedures*; and the *Standards for Educational and Psychological Testing*.

About Cooperative Personnel Services (CPS)

CPS has provided testing and human resource management services to public agencies in California for more than 60 years and to agencies throughout the U.S. and Canada for more than 14 years. Founded in 1935 as a local government services unit of the California State Personnel Board, CPS became a joint-powers agency in 1985. Today, as a self-supporting public agency, CPS offers a full range of human resource and related management consulting services to government and nonprofit agencies in the U.S. and Canada, with particular emphasis on occupational testing. CPS currently has more than 1,000 clients throughout the U.S. and Canada, distributes more than 100,000 tests per year, and administers another 150,000 tests per year. Most of CPS's clients are state or local agencies who work with CPS testing consultants to conduct occupational analysis studies, develop, produce, and administer tests, and perform validation studies for employment and licensing/certification testing. CPS has developed and gathered validity evidence for tests for numerous trade occupations such as electrician, plumber, electrical and plumbing inspector, building inspector, and for a range of non-trade occupations including police and fire occupations. Other clients work with CPS in the development and management of large licensure, certification, and competency testing programs for Paramedics, Doctors of Chiropractic, Notary Publics, Qualified Medical Evaluators, Court Interpreters, Radiologists, and Environmental Health Specialists.

Association of Firearm and Toolmark Examiners
Report of Occupational Analysis and
Certification Examination Development

Table of Contents

Introduction	i
About Cooperative Personnel Services (CPS)	i
AFTE Occupational Analysis	1
Development of the AFTE Job Analysis Questionnaire	1
Results of the AFTE Job Analysis Questionnaire	3
KSA Factor Analysis	4
Linking the Significant KSA Factors to the Significant Task Statements	4
AFTE Factor Analysis Results	5
AFTE Examination Plan Development	6
AFTE Written Examination Plans	6
AFTE Practical Examination Plans	8
AFTE Written Test Development	9
Written Test Development Training	9
Written Test Item Development Procedures	9
AFTE Item Review/Angoff Committee	10
Selection of Items for the AFTE Written Tests	11
Passpoint Setting Methodology	12
Pilot Test Administration of the AFTE Written Examinations	12
Table 1: Results of the AFTE Pilot Written Tests	13
Development of Future AFTE Written Examination Items	14
AFTE Practical Test Development	16
Linkage Between the KSAs and Practical Examination Content	17
Pilot Testing and Evaluation of the AFTE Practical Examinations	17
Examination Security	19
Content Validity of the AFTE Certification Examinations	20

AFTE Appeals Procedures	21
Appeals to the AFTE Written Examinations	21
Appeals to the AFTE Practical Examinations	21
Retention of Certification Examination Records	22
 Certification of AFTE Members	 23
 Conclusions	 24
 References	 25

List of Appendices

- Appendix A:** Job Description of the Firearm and Toolmark Examiner Classification
- Appendix B:** AFTE Job Analysis Questionnaire
- Appendix C:** Results of the AFTE Job Analysis Questionnaire
- Appendix D:** KSA Factor Analysis Groupings
- Appendix E:** Task Statement/KSA Linkages
- Appendix F:** KSAs Linked to the AFTE Written Examinations
- Appendix G:** Angoff Rating Instructions
- Appendix H:** KSA and Angoff Ratings for the AFTE Written Examination Items
- Appendix I:** Written Examination Proctor Instructions
- Appendix J:** Written Examination Comment Form
- Appendix K:** KSAs Assessed by the AFTE Practical Examinations
- Appendix L:** Practical Examination Comment Forms
- Appendix M:** AFTE Test Security Agreement Forms
- Appendix N:** Content Analysis of the AFTE Written Examinations

AFTE Occupational Analysis

The primary purpose for conducting an occupational analysis is to determine the job content areas that should be represented in the content of the certification examination. Generally, certification examinations are used to ensure that the candidates have a mastery of the various knowledge, skills, and abilities (KSAs) associated with their profession. When developing a certification examination, it is best to use a content validation strategy. By using a content validation strategy, the Test Developer can determine the domain of significant KSAs needed to successfully perform the job. By adequately sampling this domain, the Test Developer can develop assessment instruments that ensure that the passing candidates have achieved mastery of the significant KSAs.

The first step in the content validation strategy is to develop a comprehensive inventory of the job content (i.e., the tasks performed on the job). The second step is to define those KSAs that a person must possess to successfully perform each of the job duties. It is these KSAs that will comprise the content of the assessment instruments. In the occupational analysis process, each of the job tasks and KSAs are rated by Subject Matter Experts (SMEs) to determine their significance towards successful performance on the job. After determining the significant KSAs, the Test Developer must ensure that the KSAs are adequately represented in the content of the assessment instruments before the instruments will be considered content valid.

In addition, for an assessment instrument to be considered content valid, the Test Developer must demonstrate the relationship between the test content and the content of the actual job. This is accomplished through a series of “linkages.” The first step is to demonstrate the relationship between the content of the job (i.e., the job tasks) and the KSAs needed to perform each task. The second step is to demonstrate the relationship between the KSAs and the content of the assessment instrument. Thus, the job content is linked to the KSAs, and the KSAs are then linked to the test content. Hence, the purpose of conducting a job analysis is to: (1) determine the content areas that must be assessed through the examination process; and (2) to document the job-relatedness of the resulting assessment instrument. The procedures CPS used to ensure the content validity of the AFTE certification examinations are described below.

Development of the AFTE Job Analysis Questionnaire

CPS Consultants reviewed existing occupational analysis materials to compile a preliminary list of AFTE Task Statements and KSA Statements. Source material included: (1) the draft Job Description for the Firearm and Toolmark Examiner classification, as supplied by the AFTE Certification Committee; (2) drafts of critical task statements and KSA statements, as supplied by the AFTE Certification Committee; (3) the 1988 Criminalist Occupational Analysis Report, as prepared by the California Association of Criminalists and Cooperative Personnel Services; and (4) the May 6, 1994 report detailing the results of Firearm and Toolmark Examiner Occupational Analysis Survey, compiled by the California Association of Criminalists. A preliminary Job Description for the Firearm

and Toolmark Examiner classification was then developed by CPS Consultants to ensure that all of the major content areas were included in the study. This preliminary Job Description is included in Appendix A of this report. The Job Description of the Firearm and Toolmark Examiner classification was developed through an extensive review and editing process by CPS Consultants and the AFTE Certification Committee (to obtain a list of the AFTE Certification Committee members, the reader is encouraged to contact the chairperson of the AFTE Certification Committee).

Additional occupational materials were compiled by CPS Consultants for the purpose of developing comprehensive lists of Task Statements and KSA Statements. These lists were then reviewed and edited by AFTE SMEs. Many of the task and KSA statements were combined, some were deleted, and many more were added. The finalized task and KSA statements were then formatted by CPS into the AFTE Job Analysis Questionnaire. The AFTE Job Analysis Questionnaire is included as Appendix B of this report. CPS also developed customized scannable item response sheets on which the SMEs could record their responses to the AFTE Job Analysis Questionnaire.

The AFTE Job Analysis Questionnaire contained 61 Task Statements. The Task Statements were rated on two separate scales: (1) Frequency of Performance, and (2) Importance to Successful Performance of the Job. The following rating scales were used:

FREQUENCY: Rate each Task Statement on the average frequency the task was performed during the last 12 months.

- A** This task was **NOT PERFORMED** during the last 12 months.
- B** This task was performed **LESS THAN ONCE EACH MONTH** during the last 12 months.
- C** This task was performed **AT LEAST ONCE EACH MONTH** during the last 12 months.
- D** This task was performed **AT LEAST ONCE EACH WEEK** during the last 12 months.
- E** This task was performed **AT LEAST ONCE EACH DAY** during the last 12 months.

IMPORTANCE: Rate each task for how important or critical it is to the overall performance of the job. Do not consider how often you perform this task. Rate the importance of each task even if you have not performed the task during the past 12 months.

- 1** This task has **NO IMPACT** on overall job performance.
- 2** This task has a **MODERATE IMPACT** on overall job performance.
- 3** This task has a **HIGH IMPACT** on and is **ESSENTIAL** to overall job performance.
- 4** This task has an **EXTREME IMPACT** on and is **CRITICAL** to overall job performance.

The AFTE Job Analysis Questionnaire also contained 153 Knowledge Statements, 53 Skills Statements, and 106 Ability Statements. The KSA Statements were rated on two separate scales: (1) Importance to Successful Performance of the Job, and (2) the degree of competency in this KSA that is Expected at Certification. The actual rating scales used were as follows:

IMPORTANCE: Rate each KSA Statement in relation to its importance for successfully performing the job duties of a Firearm and Toolmark Examiner.

- A** This KSA is **NOT NEEDED** to successfully perform the job.
- B** This KSA is **SLIGHTLY IMPORTANT** for successfully performing the job.
- C** This KSA is **IMPORTANT** for successfully performing the job.
- D** This KSA is **VERY IMPORTANT** for successfully performing the job.
- E** This KSA is **CRITICALLY IMPORTANT** for successfully performing the job.

COMPETENCY LEVEL EXPECTED AT CERTIFICATION: Rate each KSA Statement as to what level of competency you believe is needed at the time of initial certification to perform the Firearm and Toolmark Examiner job duties.

- 1** Competency in this KSA is **NOT ESSENTIAL** at the time of certification.
- 2** **SOME COMPETENCY** in this KSA is **ESSENTIAL** at the time of certification.
- 3** **A HIGH DEGREE OF COMPETENCY** in this KSA is **ESSENTIAL** at the time of certification.
- 4** **FULL COMPETENCE** in this KSA is **ESSENTIAL** at the time of certification.

Results of the AFTE Job Analysis Questionnaire

The AFTE Job Analysis Questionnaire was sent out by AFTE to all 643 of its members. The membership was allotted approximately three weeks in which to complete the workbooks and mail the response sheets back to CPS. A total of 203 usable response sheets (31.6 percent of the entire AFTE membership) were received by CPS (10 of the response sheets received by CPS were not complete and, thereby, eliminated from the analysis). Those rating scales that were represented by alphabetical characters were quantified for the analysis (A=1, B=2, C=3, D=4, E=5). The results of the AFTE Job Analysis Questionnaire (sample sizes, mean ratings and standard deviations) are included in Appendix C of this report.

Means and standard deviations were calculated for each of the Task and KSA rating scales. The standard deviations of the ratings were examined to identify those rating scales with distributions that did not approximate a normal curve distribution (e.g., bi-modal score distributions). Any “not-normal” rating distributions would have required special analysis to ascertain the reasons why the rating distribution was not normal. However, no unusual rating distributions were found among the Task or KSA ratings.

The mean scores of the task statement frequency ratings and importance ratings were examined. All tasks that did not receive a mean score of at least 2.0 on the frequency scale and a mean score of 2.0 on the importance scale were considered non-significant, and eliminated from further analytical considerations. A score of 2.0 (or less) on the frequency scale meant the task was performed less than once a month on average, and a score of 2.0 (or less) on the importance scale meant the task had less than a moderate impact on overall job performance. Of the 61 task statements rated in the job analysis questionnaire, 14 were eliminated, and 47 were retained for further analysis.

The mean scores of the KSA ratings were examined. All KSA statements that did not receive a mean score of at least 2.0 on the Importance scale and a mean score of 2.0 on the Competency Level Expected at Certification scale were considered non-significant, and eliminated from future analytical considerations. A score of less than 2.0 on the Importance scale meant the KSA was less than slightly important for successful performance of the job. A score of less than 2.0 on the Competency Level Expected at Certification scale meant that only a small amount of competency in that KSA was needed by the worker at the time of certification. Of the 153 Knowledge Statements rated in the job analysis questionnaire, 29 were eliminated, and 124 were retained for further analysis. Of the 53 Skills Statements rated in the job analysis questionnaire, 9 were eliminated, and 44 were retained for further analysis. Of the 106 Ability Statements rated in the job analysis questionnaire, 0 were eliminated, and all 106 were retained for further analysis.

KSA Factor Analysis

A factor analysis procedure was then performed on the significant KSA statements using the scores from the Importance Rating Scale. A factor analysis is a statistical procedure that examines the intercorrelations among the items on a particular rating scale. Items that are highly correlated with each other are combined into groups. CPS Consultants examined the groups of items to determine the commonalities in their content that would cause the highly similar ratings. From the factor analysis, 14 Knowledge factors were identified, 8 Skills factors were identified, and 14 Ability factors were identified (listed on Page 6). (Note: the order of the KSA factors listed above is based on the strength of the intercorrelations between the items in each factor. The order does not represent the relative importance of the factors). The KSA statements which loaded onto each factor are listed in Appendix D.

Linking the Significant KSA Factors to the Significant Task Statements

On April 14th and 15th, 1998, eight AFTE Subject Matter Experts (all members of the AFTE Certification Committee) met with CPS staff in Sacramento, California for the purpose of linking the significant task statements to the KSA factor groupings. The linkage procedure involves examining each task statement and coming to a group consensus on whether each individual KSA factor is needed to successfully perform that task. This linkage is necessary for the documentation of the job-relatedness of the resulting written and performance examinations. Because of the success of the factor analysis procedure (described above), CPS Consultants had the AFTE SMEs link the significant task statements with the KSA factors (as opposed to linking the tasks to the individual KSA statements). A total of 1,692 possible KSA factor-Task Statement linkages were examined by the eight-member Subject Matter Expert panel. (Note: if the individual KSAs were used in the linkage process, then the SMEs would have been required to perform 13,348 linkages). The results of the Task-KSA linkages are included in Appendix E of this report.

AFTE Factor Analysis Results

Knowledge Factors

- Factor 1: Laboratory Equipment
- Factor 2: Firearms/Ammunition
- Factor 3: Collection and Preservation of Evidence
- Factor 4: Ballistics
- Factor 5: Safe Work Practices
- Factor 6: Distance Determination
- Factor 7: Firearm Design and Identification
- Factor 8: Documentation of Physical Evidence and Analytical Results
- Factor 9: Professional Development
- Factor 10: Toolmark Identification
- Factor 11: Firearm-Related Tools
- Factor 12: Examination of Evidence
- Factor 13: Maintaining Professional Ethical Standards of Conduct
- Factor 14: Tools/Equipment

Skills Factors

- Factor 1: Collection of Evidence
- Factor 2: Serial Number Restoration Techniques
- Factor 3: Firearms
- Factor 4: Firearm-Related Tools
- Factor 5: Gunshot Residues
- Factor 6: Firearm Safety
- Factor 7: Laboratory Techniques
- Factor 8: Laboratory Equipment

Abilities Factors

- Factor 1: Ballistics
- Factor 2: Firearm/Ammunition Analysis
- Factor 3: Examination and Documentation of Evidence
- Factor 4: Gunshot Residues
- Factor 5: Ammunition Components
- Factor 6: Firearm Design
- Factor 7: Collection of Evidence
- Factor 8: Matters and Procedures of Evidence Examination
- Factor 9: Dissemination of Evidence and Analytical Results
- Factor 10: Safe Work Practices
- Factor 11: Testifying in Legal Proceedings
- Factor 12: Professional Development
- Factor 13: Toolmark Identification
- Factor 14: Preservation of Evidence

AFTE Examination Plan Development

The primary purpose of the Examination Plan is to specify the content of the examination. To ensure that an assessment instrument is content valid, the Test Developer must develop an Examination Plan which documents that the significant KSAs are adequately sampled by the assessment instrument. Qualitatively this refers to the identification of those KSAs that will be assessed by the certification examination. Quantitatively, the Examination Plan must specify the percentage of the examination's content (i.e., the weights) which will assess each specific KSA area.

AFTE Written Examination Plans

To develop the examination plans for the AFTE written examinations, CPS Consultants utilized the results of the KSA factor analyses. CPS Consultants reviewed the factors, and combined those with similar content; the result of which were the following eight examination content areas listed below (in parenthesis are the original KSA factors that comprise each content area). Note: skills are not typically measured in a written examination, and have not been included in this section. The candidates' skills will be assessed through the practical examinations.

- 1. Collection and Preservation of Evidence**
(Knowledge Factor 3, and Ability Factors 7 and 14)
- 2. Examination of Firearm Evidence:** includes: Firearm/Ammunition, Firearm Design and Identification, Laboratory Tools and Equipment, and Ballistics
(Knowledge Factors 1, 2, 4, 7, 11, 12, 14, and Ability Factors 1, 2, 3, 5, and 6)
- 3. Examination of Gunshot Residue Evidence:** GSR and Distance Determination
(Knowledge Factors 6 and 12, and Ability Factors 3 and 4)
- 4. Examination of Toolmark Evidence**
(Knowledge Factors 10 and 12, and Ability Factors 3 and 13)
- 5. Documentation of Physical Evidence & Analytical Results**
(Knowledge Factor 3, and Ability Factor 9)
- 6. Safe Work Practices**
(Knowledge Factor 5, Skill Factor 6, and Ability Factor 10)
- 7. Professional Development**
(Knowledge Factor 9, and Ability Factors 11 and 12)
- 8. Maintaining Professional Ethical Standards of Conduct**
(Knowledge Factor 13)

In April of 1998, the eight AFTE Subject Matter Experts from the AFTE Certification Committee reviewed all of the significant KSA statements to determine which assessment instrument(s), (i.e., the written examinations and/or practical examinations), would best assess each KSA. The result of this process was the development of a list of KSAs for each of the three AFTE written examination competency areas. CPS Test Development Consultants then reviewed this list, and eliminated those KSAs that would be difficult to assess in a multiple choice examination format (e.g., ability to

recognize bullet impact sites at crime scene, and the ability to identify a cast bullet as having come from a particular bullet mold). The KSAs linked to each AFTE written examination are included in Appendix F of this report. Most of the KSAs were included on more than one of the examination lists.

Through this linkage process, CPS consultants were able to determine the number of KSAs that are associated with each of the examination content areas for the AFTE written examinations. CPS Consultants then divided the number of KSAs in each examination's content area by the total number of KSAs that were linked to that particular examination, the result of which was the weight assigned to that content area. This process treats all of the significant KSAs equally by assigning them equal representation in the examinations. Such a procedure is commonly used in mastery testing situations (e.g., certification examinations), where the goal is to assess the widest range of KSAs as possible. The AFTE written examination plans are listed below (note: the percentages for each of the content areas have been rounded to the nearest whole number):

AFTE Firearm Written Examination

Collection and Preservation of Evidence: (13 %)
Examination of Firearm Evidence: (60 %)
Examination of Gunshot Residue Evidence: (6 %)
Examination of Toolmark Evidence: (4 %)
Documentation of Physical Evidence and Analytical Results: (4 %)
Safe Work Practices: (7 %)
Professional Development: (4 %)
Maintaining Professional Ethical Standards of Conduct: (2 %)

AFTE Gunshot Residue Written Examination

Collection and Preservation of Evidence: (16 %)
Examination of Firearm Evidence: (10 %)
Examination of Gunshot Residue Evidence: (39 %)
Examination of Toolmark Evidence: (0 %)
Documentation of Physical Evidence and Analytical Results: (5 %)
Safe Work Practices: (9 %)
Professional Development: (6 %)
Maintaining Professional Ethical Standards of Conduct: (3 %)

AFTE Toolmark Written Examination

Collection and Preservation of Evidence: (21 %)
Examination of Firearm Evidence: (14 %)
Examination of Gunshot Residue Evidence: (0 %)
Examination of Toolmark Evidence: (39 %)
Documentation of Physical Evidence and Analytical Results: (6 %)
Safe Work Practices: (8 %)
Professional Development: (9 %)
Maintaining Professional Ethical Standards of Conduct: (3 %)

CPS Consultants recommend that any new AFTE written examinations include a minimum of 100 multiple choice test items. The inclusion of at least 100 items will allow for adequate sampling of all the significant KSAs, and it will increase the probability that the overall test statistics will yield professionally-acceptable results. The percentages of test questions assessing each content area have been rounded off to sum to 100 percent. CPS recommends that the actual items included in future AFTE written examinations reflect the above percentages as much as possible. Note: While the above-listed percentages represent the ideal make-up of the AFTE written examinations, the actual number of items in the AFTE written examination content areas will be largely dependent upon the number of usable written test items in each of the competency areas developed by AFTE. Additionally, there exists some overlap in the content of the AFTE competency areas (e.g., many of the items refer to more than one KSA) which will have an effect upon the final composition of the AFTE written examinations.

AFTE Practical Examination Plans

The examination plans for the AFTE practical examinations were developed by the AFTE Certification Committee in June of 1998 in a meeting held in Sacramento at CPS. During this meeting, the Certification Committee determined the content, format, scoring dimensions, and administration procedures for the three AFTE practical examinations. It was the consensus of the group that the practical examinations should reflect the type of work activities that are most commonly performed on the job. By determining the content of the AFTE practical examinations, the AFTE Certification Committee was, in essence, also developing the examination plans for the AFTE practical examinations.

AFTE Written Test Development

The KSAs were ranked by the AFTE Certification Committee according to the ratings they received through the AFTE Job Analysis Questionnaire. These ranked lists were then distributed to the item writers with instructions as to which KSAs to emphasize when researching the job-related materials for potential test questions. The goal of this procedure was to ensure that the most significant KSAs were represented in the resulting AFTE written and performance examination, with as little overlap between the KSAs as possible.

Written Test Development Training

On April 15th, 1998, the eight Subject Matter Experts from the AFTE Certification Committee participated in a half-day workshop on *Developing Multiple Choice Test Questions*. The workshop was facilitated by CPS Test Development Consultants and was geared towards assisting the participants in training other AFTE member SMEs to write multiple choice test questions for the three AFTE competency areas. Workshop topics included: how to develop various types of multiple choice items; developing the various components of multiple choice items; item sensitivity review procedures; item writing clues and errors; test development validity issues; item writing strategies; and how to properly edit multiple choice items. During the workshop, participants were administered a number of item writing and item editing exercises. All participants were given copies of CPS' *ABCs of Item Writing and Review* to distribute to their item writers.

Written Test Item Development Procedures

All AFTE written test items were developed using the four-alternative multiple-choice format. This format was chosen because it: (1) is one of the best formats for differentiating between good and poor test candidates; (2) affords a lot of statistical power for analysis; (3) is highly recognizable and readily acceptable by most test candidates; (4) can be objectively scored, and machine scored; and (5) has a lower chance score (25 %) than other types of written item formats (such as true/false or three-alternative multiple choice questions).

Item Writers were recruited by the AFTE Certification Committee members. Over 80 AFTE members were asked to participate in the item writing process. When selecting item writers, AFTE attempted to obtain SMEs with a wide range of experience, as well as geographical representation. Each potential AFTE item writer was provided with a copy of CPS' *"The ABCs of Item Writing and Review,"* as well as given individualized instruction (for those SMEs who requested it). Item writers were each assigned a limited number of specific KSAs on which to develop multiple choice items. This limited the amount of knowledge (or potential advantage) any one SME could have regarding the AFTE written examinations.

For each test item, SMEs were required to document: (1) a reference that could be used to verify the correct response, and (2) the KSAs that the test item assessed. Items written by the AFTE SMEs were mailed directly to CPS, where they were reviewed and edited by CPS Test Development Consultants. This procedure ensured that no AFTE members (i.e., potential examination candidates) had access to items other than the items they had personally written. SMEs were allowed to keep no copies (either on paper or in electronic format) of the items they developed. Item writers were asked to send all examination materials (including scratch paper) to CPS, or to destroy the materials in a secure manner. CPS Test Development Consultants reviewed the items for content, clarity, form, and relevance. Items were also reviewed to minimize the number of duplicate items that were entered into the AFTE item bank. A total of 1,000 items were entered into CPS's computerized item bank. Once entered, the AFTE test items were then subjected to CPS's strict item review procedures to ensure that the items were properly formatted and entered into the computer.

AFTE Item Review/Angoff Committee

In March of 1999, a panel of twelve AFTE SMEs was assembled in Sacramento to review the AFTE written items. When selecting the panel members, AFTE attempted to obtain SMEs with a wide range of experience, as well as geographical representation. The AFTE written test items were reviewed and edited in a group discussion forum. The SMEs reviewed each item for clarity, relevance, and to ensure that the keyed response was the correct answer. The SMEs were not supplied with the correct answer when initially reviewing the test items (although this information was available upon request). Reference materials were available to the SMEs so that they could verify the keyed responses. The SMEs also reviewed the distractors to ensure: (1) they were distinctly incorrect, and (2) they were logical alternatives to the question being posed. Any problems with test items that the SMEs could not resolve were documented in the item bank (in the item "Comments" field) to alert the Test Developer not to select these items into specific tests. In the item review process, the SMEs were also asked to judge whether each item was relevant for inclusion in the Firearm, Toolmark, and/or Gunshot Residue examinations. If at least 50 percent of the SMEs marked an item as relevant for a particular examination, then that item was flagged in the computer as being eligible for inclusion into that particular examination. (Note: 64 of the 1,000 AFTE items were not reviewed because they were incomplete (e.g., missing KSAs or references).

In licensure and certification testing it is a common practice to use a criterion-referenced approach to setting the passpoint (cutoff score) for an examination. This approach provides a defensible rationale for identifying a cutoff score. The main rationale behind criterion-referenced cutoff scores is that a line must be drawn to identify those candidates who can demonstrate sufficient knowledge to be certified. To establish the passpoint for the AFTE written examinations, CPS used the Angoff method. The Angoff method defines the cutoff score as the most likely score a minimally competent applicant is likely to achieve. Candidates scoring below this level are believed to lack sufficient knowledge, skills, or abilities to be certified. For each test item, SMEs are asked to estimate the probability that a minimally-competent applicant will get the answer correct. To accomplish this, the SMEs are asked to estimate what percentage of minimally acceptable candidates would achieve a

correct response to each test item. These percentages are then averaged for all the items in a particular examination to determine the minimally acceptable score, or passing score, for that test.

The SMEs each assigned the test items a difficulty rating using the Angoff method. The Test Review Committee members were provided formal training on how to assign the Angoff ratings. The Angoff Rating Instructions are included in Appendix G. Using the Angoff method, SMEs were asked to estimate how many of the “Minimally-Acceptable Candidates” (out of a group of 100) would be expected to obtain the correct response for each test question. According to AFTE, a Minimally-Acceptable Candidate is one who has completed the applicable training in the occupation and possesses a minimum of five years of experience in his/her discipline. SMEs were instructed to assign each item an Angoff rating between 25 (representing the chance score of a four-alternative multiple choice item) and 100 (indicating that all test candidates should be expected to get the item correct). AFTE SMEs were required to assign an Angoff rating to each item (independent of the other raters).

After the SMEs rated the first AFTE item, the CPS Consultant conducted a group discussion on the criteria used by each SME to rate the item. Through this discussion, the CPS Consultant verified that each of the Panel Members understood the rating process and was correctly marking their responses on the Angoff Response Sheet. Significant differences (greater than 15 points) between the SMEs on particular items were discussed in a group forum. The purpose of these discussions was to have the SMEs explain the criteria they used when rating the items, so as to ensure the other SMEs considered all pertinent aspects of the item when making their ratings. The SMEs were then given an opportunity to change their ratings based on these discussions. However, the SMEs were not required to change their ratings at any time during the rating process. This process continued for the first 25 AFTE items. The SMEs were then allowed to rate each of the AFTE items independently. Throughout the Angoff process, CPS Consultants monitored the SME ratings to identify those items with large score dispersions (i.e., over 15 points) for group discussion. The Angoff ratings were averaged over the 12 SMEs to compute an overall Angoff rating for each item.

Selection of Items for the AFTE Written Tests

Two versions of the Firearm, Toolmark, and Gunshot Residue written examinations were assembled by CPS Test Development Consultants (a total of six written examinations, in all). Each test version consisted of 100 items, 20 of which were classified as “anchor items” and included in both versions of the examination. Anchor items allow the CPS Test Consultant to “equate” the examinations to determine if they are of equal difficulty. (Note: the equating procedure is performed after the tests versions have acquired a large number of statistical history and, as such, will not be examined in this report). Therefore, a total of 180 items were selected for each of the three examination competency areas. There was some item overlap between the three competency areas in those areas pertinent to all three examinations (e.g., collection and preservation of evidence, ethical conduct, and testifying in court).

A number of different criteria were used by the CPS Consultant to select the items into the test

versions. First, the item KSAs were reviewed to ensure that a high proportion of the significant KSAs (as identified in the examination plans) were represented in the written tests. The CPS Consultant also reviewed the item content to minimize the number of items with similar content being included in the same examination version. The CPS Consultant also reviewed the item Angoff ratings to ensure that alternate versions of the same test were of equal difficulty. The KSAs and Angoff ratings assigned to those items selected for the AFTE written examinations are included in Appendix H. As much as possible, the CPS Consultant also attempted to balance the Angoff ratings between the three different test content areas. By doing so, it would allow the CPS Consultant to set similar passpoints for all three of the AFTE written examinations. The average Angoff ratings for the AFTE Written Tests were as follows:

<u>Test</u>	<u>Version</u>	<u>Average Angoff Rating</u>
Firearm	A	71.26
Firearm	B	71.30
Gunshot Residue	A	71.26
Gunshot Residue	B	71.26
Toolmark	A	71.98
Toolmark	B	71.98

Passpoint Setting Methodology

CPS Consultants recommended that the passpoints for the AFTE written examinations be set at 70. The CPS Consultants felt that the passpoints should be set slightly lower than the Angoff rating in order to compensate for any rater error that may have occurred during the Angoff rating process. Additionally, CPS Consultants believed that the 70% passpoints are appropriate for the AFTE written examinations, given that they are only one component in the AFTE certification process (the practical test being the other).

Pilot Test Administration of the AFTE Written Examinations

The AFTE written examinations were administered to various AFTE members at the annual AFTE Training Seminar, in July of 1999. The written examinations were administered by members of the AFTE Certification Committee. The proctor instructions describing how the AFTE written examinations were administered are included in Appendix I. After completing an AFTE written examination, candidates were able to record their opinions about the examination questions on the “Examination Comment Form.” A copy of the Examination Comment Form is included in Appendix J.

The written examination materials were then sent to CPS for scoring and analysis. CPS Consultants scored the AFTE written pilot tests, and performed statistical item analysis (computing: p-values; bi-serial correlations; and distractor analyses) to identify miskeys and malfunctioning distractors. Based

on the statistical analyses, CPS identified those items that were either potential “miskeys” or had more than one correct response. These items, as well as the item comments that were made by the candidates on the “Examination Comment Form,” were then reviewed by select members of the AFTE Certification Committee. Upon review, the members of the Certification Committee made their recommendations of item edits to CPS. While performing this review, the AFTE reviewers did not have access to any of the individual candidate scores while performing their edits. The written examinations were then rescored. The test results of the AFTE pilot written examinations are included in Table 1.

Table 1
Results of the AFTE Pilot Written Examinations

Test	Version	N	Mean	SD ¹	Reliability	SEM ²
Firearm	A	16	79.38	9.75	0.88	3.38
Firearm	B	15	83.07	7.32	0.81	3.19
GSR	A	12	87.50	6.57	0.82	2.81
GSR	B	10	85.00	8.68	0.89	2.83
Toolmark	A	9	84.33	6.30	0.79	2.89
Toolmark	B	8	85.12	4.94	0.65	2.91

Upon reviewing the AFTE written pilot test results, the CPS Consultants’ greatest concern is the that the mean ratings (or difficulty levels) indicate the examinations may be too easy. The mean ratings for most of the written tests were an average of 12.56 points higher than the Angoff ratings for each examination. However, the sample sizes are far too low to make any generalizations regarding how the tests will ultimately perform. The elevated mean scores could be an indication that those candidates who participated in the pilot tests were, on the whole, some of the more knowledgeable AFTE members. The item edits made by the members of the AFTE Certification Committee (described above) as well as edits made prior to the administration of the written tests at the AFTE Training Seminar (to be held in June 2000), should have the effect of making the items more difficult for future test administrations. Regardless of the causes, the mean scores for the AFTE written examinations should be closely monitored to determine if the AFTE written examinations need additional edits to make them more difficult.

¹ SD = Standard Deviation. The Standard Deviation is a measure of score dispersion; higher standard deviations indicate the scores are more widely spread.

² SEM = Standard Error of Measurement. The Standard Error of Measurement is a measure of how accurate the examination is at assessing the candidate scores. Generally, small SEMs are associated with more reliable examinations.

Although the sample sizes for each of the written examinations are too low to draw any definitive conclusions, it would appear that the AFTE written examinations: (1) are reliable; and (2) that the alternate versions (Version A and Version B) of each written test are comparable. An examination's reliability coefficient is an indication of how well an examination consistently measures a particular candidate attribute (or KSA). In other words, if an examination were given today and then re-administered to the same group of candidates one year later, a reliable examination would yield similar results for both exam administrations. Reliability coefficients above .70 are considered to be an indication of a reliable examination for examinations with sample sizes greater than 100. Generally, tests with smaller sample sizes will have smaller reliability coefficients. However, as evidenced by the statistics in Table 1, the reliability coefficients for the AFTE written examinations are extremely high (especially considering the low sample sizes). The sample sizes for these examinations are too low to make any generalizations regarding the size of the examinations' standard deviations and standard error of measurements.

The mean scores and standard error of measurement (SEM) for the different written examinations appear to be comparable, even given the small sample sizes. Evaluation of the anchor items for each of the test versions also yields comparable results. These initial statistical results would suggest that the different versions of the tests are equally difficult, and might be able to be used interchangeably. Although the initial test administration results are very encouraging, larger sample sizes ($N \geq 30$) must be obtained for all test versions before any generalizations can be made regarding the written examinations' statistical soundness.

Development of Future AFTE Written Examination Items

Future AFTE written examination items should be developed using the same procedures described in this report. First, the items should be developed by qualified SMEs (i.e., those AFTE members that have already passed the certification examination for which the test item is intended). Next, the new test items should be reviewed by Test Development Consultants to ensure the items are clearly written and properly formatted. The new test items should then be reviewed by members of the AFTE Certification Committee to ensure: relevance (i.e., whether the test questions are relevant to the subject areas); clarity (i.e., whether the test questions are understandable as written); that the keyed responses are the only correct answers; and that there is minimal overlap between the content of the test questions in any particular examination version. Finally, the new test items should be reviewed and rated by a properly-trained Angoff committee of 8 - 12 AFTE members. Once the items have been assembled into new test versions, the Test Development Consultant should compute the average Angoff rating for the test version so as to recommend and set the appropriate examination passpoint. Additional suggestions for future AFTE written examinations are included on page 8 of this report.

To ensure an impartial test process, the written examinations would best be scored by an agency independent of AFTE. Additionally, the existing AFTE written tests should be statistically evaluated periodically to ensure the items, and test as a whole, are performing properly. The following item

statistics should be calculated: item p-values (difficulty analysis); and item distractor analyses (biserial and/or point-biserial correlations). Additionally, the following test statistics should also be calculated: test means; test standard deviations; and standard error of measurements. The purpose of the statistical analyses is to ensure that the written examinations are accurately and reliably assessing the appropriate KSAs.

AFTE Practical Test Development

In June of 1998, the AFTE Certification Committee convened in Sacramento to discuss the content, format, and administration procedures for the three AFTE Practical Tests. It was determined by the Certification Committee that two alternate versions of each of the practical examinations should be developed, with each version having two separate components. After much discussion, the AFTE Certification Committee decided that the AFTE Practical Examinations would consist of the following test components:

Firearm Practical Examination:

Version 1:

Component 1: 9mm Bullet Comparison

Component 2: 9mm Cartridge Case Comparison

Version 2:

Component 1: 45 ACP Bullet Comparison

Component 2: 45 ACP Cartridge Case Comparison

Gunshot Residue Practical Examination:

Version 1:

Component 1: 9mm Distance Determination

Component 2: 12 Gauge Shot Dispersion Comparison

Version 2:

Component 1: 45 ACP Distance Determination

Component 2: 20 Gauge Shot Dispersion Comparison

Toolmark Practical Examination:

Version 1:

Component 1: Diagonal Plier Striated Mark Comparison

Component 2: Number Stamp Impression Comparison

Version 2:

Component 1: Nipper Plier Striated Mark Comparison

Component 2: Plier Jaw Impression Comparison

For each test component, the AFTE Certification Committee also determined: the numbers and types of physical materials needed (e.g., unknowns, standards, and shot dispersion patterns); the test instructions and administration procedures; the suggested time limits; the scoring criteria; and the Certification Committee member who would develop the physical components. The Committee also developed an appeals process by which the test candidates can (and cannot) appeal the results of their Practical Examinations.

Linkage Between the KSAs and Practical Examination Content

The Certification Committee member who developed the physical materials for a particular test component was also asked to determine which significant KSAs were assessed by the test component. As stated earlier, these linkages are needed to demonstrate the relationship between the test content and the actual duties performed on the job, thus, demonstrating the “job-relatedness” and content validity of the AFTE Practical Examinations. The linkages between the AFTE Practical Examination test components and the KSAs are included in Appendix K.

Pilot Testing and Evaluation of the AFTE Practical Examinations

Prior to arriving at the March 1999 Angoff meeting in Sacramento, the twelve Angoff Panel members were each sent one version of each of the three AFTE practical examinations (six were sent Version 1 of each examination, and six were sent Version 2). The SMEs were allowed to take the AFTE Practical examinations in their own laboratories, using their own equipment. Immediately after completing each AFTE practical examination, the Angoff Panel members were asked to complete an AFTE Practical Examination Comment Form for each test component. Copies of the Comment Forms are included in Appendix L. In the Comment Forms, the Angoff Panel members were asked to respond to the following questions:

Time Limits:

1. Is there too much or not enough time for a competent examiner to complete the exercise?

Test Instructions:

1. Were the test instructions clear and easy to understand?
2. Did the test instructions have all the information you needed to begin the exercise?

Examination Materials:

1. Were the examination materials clearly labeled or marked?
2. Were the examination materials representative of what is actually performed on the job?
3. Were the examination materials appropriately prepared for the exercise?
4. How could the examination materials be improved for this exercise?

Exercise Difficulty Levels:

1. Is the test exercise either too easy or too difficult? Please explain your response.

Scoring Criteria:

1. Was the scoring criteria used for this exercise clear and easy to understand?
2. Is the scoring criteria fair to all candidates?

Worksheets:

1. Were the exercise response worksheets clear and easy to understand?
2. Do the exercise response worksheets allow the candidate to accurately and thoroughly record his/her responses?

Packaging:

1. Were the examination materials appropriately packaged for this exercise?
2. Were the packaging requirements appropriate for this examination?

Proctors:

1. Are proctors needed to administer this examination exercise?

Overall:

1. Is this exercise assessing the appropriate job knowledge, skills, and abilities?

When the Angoff Panel convened in Sacramento, the SMEs spent a full day discussing their responses to the above-listed questions. Many minor modifications to the AFTE Practical Examinations were proposed, such as: packaging the evidence in resealable bags; changes to the Examination Instructions and Candidate Response Sheets; and suggestions on how to best mark the evidence. The Angoff Panel also determined that it would be beneficial to have a proctor administer the AFTE Practical Examinations (although exact details of this have not been determined at this time). Some of the more significant recommendations by the Angoff Panel members included: developing a more distinct set of shot shell dispersion patterns; and ways to make the Toolmark Practical Examinations more difficult (the Angoff Panel felt that the toolmark test components were too easy). The overwhelming consensus of the AFTE Item Review Committee was that the AFTE Practical Examinations were highly job-related, and that they accurately assessed the appropriate KSAs. The AFTE Practical Examinations will be modified according to the recommendations made by the Angoff Panel members. The completed changes to the AFTE Practical Examinations will then be reviewed by a post-Angoff committee (consisting of one member from the Certification Committee and one member of the original Angoff Panel) to ensure that the modifications reflect the KSAs they are intended to assess, and accurately reflect the opinions of the Angoff Panel.

Examination Security

From the onset of this test development project, the Association of Firearm and Toolmark Examiners Certification Committee has taken very stringent precautions to ensure examination security. AFTE's primary examination security concern was the fact that candidates that were writing some of the examination items would actually be taking the AFTE certification examinations. To counteract this, the AFTE item writers were each assigned a limited number of specific KSAs on which to develop multiple-choice items. This limited the amount of knowledge (or potential advantage) any one SME could have regarding the AFTE written examinations. Items written by the AFTE SMEs were mailed directly to CPS, where they were reviewed and edited by CPS Test Development Consultants. This procedure ensured that no AFTE members (i.e., potential examination candidates) had access to items other than the items they had personally written. SMEs were allowed to keep no copies (either on paper or in electronic format) of the items they developed. Item writers were asked to send all examination materials (including scratch paper) to CPS, or to destroy the materials in a secure manner.

Test Security Agreements were obtained from everyone associated with the AFTE written and practical examination development processes, including: members of the AFTE Certification Committee, Angoff Panel members, and written examination item writers. Test Security Agreements were also obtained during the test administration process from examination proctors and test candidates. Copies of the AFTE Test Security Agreements are included in Appendix M. Violations of the AFTE Test Security Agreements could result in various disciplinary actions, including expulsion from AFTE.

Content Validity of the AFTE Certification Examinations

For an assessment instrument to be considered content valid, the Test Developer must demonstrate the relationship between the test content and the content of the actual job. For the AFTE Certification Examinations, this was accomplished through a series of “linkages.” CPS Consultants demonstrated the relationship (or “linkage”) between the content of the job (i.e., the job tasks) and the KSAs needed to perform those tasks (see Appendix E). CPS Consultants then demonstrated the relationship between the KSAs and the content of the AFTE written examinations (see Appendix H) and the AFTE practical examinations (see Appendix K). All of the items included in the AFTE written examinations were rated “as appropriate for that particular examination” by a majority of the members of the Angoff Committee. The content of the AFTE written examinations (as expressed in percentages of items linked to the examination content areas) is summarized in Appendix N. CPS Consultants feel that the content of the current AFTE written examinations (Versions A and B) are sufficiently similar to the proposed percentages expressed in the examination plans, as to attest to the examinations’ content validity. Passpoints for the AFTE written examination were set using the Angoff method for determining the “minimally-qualified” candidate, based on the AFTE standard of five-years of examiner experience.

As with the vast majority of certification examination, the AFTE Certification Examinations were not designed to maximally differentiate between the candidates. Certification and licensing examinations are generally used to ensure that the candidates have achieved a minimum competency level of the various knowledge, skills, and abilities associated with their profession. By using a content validation strategy, CPS Test Development Consultants have determined the domain of significant KSAs needed to successfully perform the job. By adequately sampling this domain, CPS Consultants have developed assessment instruments that ensure that the passing candidates have achieved mastery of the significant KSAs.

AFTE Appeals Procedures

AFTE will convene an Appeals Panel whose sole purpose will be to review any appeals to the AFTE written and practical examination components. Some of the specific appeal procedures are documented below:

Appeals to the AFTE Written Examinations

Applicants who failed any written examination will have thirty (30) calendar days from the date the results are received by the applicant to appeal the results. All of the questions on the AFTE written examinations have been reviewed by a panel of AFTE members and each has been determined to be properly worded, applicable (relevant) and the correct answer to each has been confirmed. Therefore, no appeal as to the clarity, applicability (relevancy) or correctness of the keyed response will be allowed. The Appeal Panel will, however, review answer sheets for entry errors. All notifications will be by certified mail, return receipt requested. The Chairman of the AFTE Appeals Panel will review the applicant's scoring sheet (not the examination booklet) and re-score the exam. Any discrepancies will be noted and confirmed by another member of the Appeal Panel. If the scoring changes are in the applicant's favor and the new score puts the applicant above the pass/fail score, it will be noted and the applicant will be notified that the written exam will be considered passed. Otherwise, the applicant will be notified that the appeal failed to change the outcome of the original scoring. The appeal process will take no longer than thirty (30) calendar days from the date the letter of appeal is delivered to the Chairman. The decision of the Appeal Panel is final.

Appeals to the AFTE Practical Examinations

The appeal of a practical examination must also be directed in writing to the Chair of the Appeal Panel within thirty (30) calendar days of the date the applicant was notified of the results. The Chairman will first review any notes, sketches and/or photographs made by the applicant during the practical exam. If there are photographs, the Chairman will attempt to verify that the specimens depicted in the photos are in fact the specimens provided to the applicant for the test and examined during the test. The Chairman will independently evaluate the correctness of the applicant's responses that are in dispute by reexamining the unknown and standard (known) specimens. The Chairman will then forward the examination package to a second member of the Appeal Panel for an independent review to be conducted in the same fashion. The second member will not be the person who originally prepared the exam. The examination package will then be sent to the member of the Certification Committee who prepared that portion of the test. That person will review and evaluate the entire package for a third independent review. If at least two of the three reviewers agree that there was sufficient basis for the conclusion(s) reached by the applicant during the examination, the applicant's answer(s) will be considered correct. If the result of this review process puts the applicant above the pass/fail score, the applicant will be notified that he/she passed that portion of the exam. Otherwise, the applicant will be notified that the appeal process failed to change the outcome of the original scoring. This notification will be by certified mail, return receipt requested.

The appeal process will take no longer than 60 calendar days from the date the appeal letter is delivered to the Chairman. The decision of the Appeal Panel is final.

Retention of Certification Examination Records

It will be the policy of AFTE to destroy all candidate test materials (worksheets, photos, etc.) for all those who pass an examination. For those applicants who fail, their test materials (worksheets, photos, etc.) will be retained only for thirty calendar days from the date the applicants have received notices of their results unless an applicant files an appeal. For those applicants who appeal, their test materials will be kept for the duration of the appeal. After the appeal and review process is final, the review records and candidate test materials (worksheets, photos, etc.) will be destroyed but a notation will be made that an appeal was made, the date of the appeal, and the results of the appeal. AFTE will only keep a record of an applicant's application(s), date of testing, the number and/or letter designation of the test version/component administered, whether or not the applicant passed, and the date and result of any appeal. For those who failed an examination, all records of their failing will be removed once the applicant has re-tested and passed.

Note: the AFTE Examination Appeals Procedures and Retention of Certification Examination Records procedures described above were developed in this draft form (6/11/99) by the AFTE Certification Committee with input from CPS Consultants. The Examination Appeals Procedures and Retention of Certification Examination Records procedures are currently pending approval by the AFTE Board of Directors. CPS neither approves or disapproves of the AFTE Examination Appeals Procedures and takes no responsibility for any legal actions that are the direct result of their use or misuse. The AFTE Examination Appeals Procedures are included in this technical report so as to provide a informational resource to the reader on the AFTE test administration process.

Certification of AFTE Members

Certification examinations are generally scored on a pass/not pass basis, and the AFTE Certification Examinations are no different. To become AFTE certified, candidates must successfully achieve a passing score on both the written and practical examination. The AFTE Board of Directors determined that candidates must pass the written examination before they would be allowed to take the associated practical examination. Should a candidate pass the written examination and fail the practical examination, the candidate would only have to retake the practical examination in order to become certified. Candidates will only be allowed to retake those examination components that they have failed after six months have elapsed from the date they were administered the original examination. Additionally, the AFTE Certification Committee has established procedures by which candidates can challenge the results of both the written and practical examinations (described above).

The AFTE certification is valid for five years from the date on the certification document issued by AFTE. As the AFTE Certification Program is currently structured, members will have to successfully pass the AFTE written and practical examinations once every five years in order to become re-certified. This provision may be amended in the future; allowing for members to become recertified by accumulating credits for job-related activities, such as: attending professional conferences and seminars, publishing journal articles, and attending approved job-related courses.

Conclusions

Cooperative Personnel Services has assisted the Association of Firearm and Toolmark Examiners in the development of a series of content-valid certification examinations in three different competency areas: firearm evidence examination and identification, gunshot residue evidence examination and identification, and toolmark evidence examination and identification. Each certification examination developed under this contract consists of a written examination and a practical examination. Two versions of the written and practical certification examinations were developed in each competency area. The primary purpose of the AFTE examinations is to certify that Examiners possess the required knowledge, skills, and abilities necessary to successfully perform essential duties within each of the competency areas.

The AFTE Certification Examinations were developed using a content validation strategy. The occupational analysis and test development procedures described in this report were developed in accordance with professionally-acceptable methodologies, and the resulting AFTE certification examinations are in compliance with applicable professional and legal standards, such as: the Civil Rights Act of 1991; the Americans with Disabilities Act (ADA); Age Discrimination Legislation; the *Uniform Guidelines for Employee Selection Procedures*; the *Standards for Educational and Psychological Testing*; and the *Principles for the Validation and Use of Personnel Selection Procedures*. The purpose of this report is to document the steps taken by CPS and AFTE to develop the content-valid certification examinations for the Association of Firearm and Toolmark Examiners.

CPS Consultants are confident that the AFTE Certification Examinations described in this report are consistent with state and federal legislation governing content-valid certification examinations, and that the AFTE Certification Examinations accurately reflect the depth and breadth of knowledge, skills, and abilities needed to successfully perform the Firearm and Toolmark Examiner job duties.

References

- American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (1985). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Psychological Association.
- Aiken, L. R. (1987). Testing with multiple choice items. *Journal of Research and Development in Education*, 20, 44-58.
- Angoff, W.H. (1971). "Scales, Norms, and Equivalent Scores." In R.L. Thorndike's *Educational Measurement*, 2nd edition, American Council on Education: Washington: D.C.
- Board, C., & Whitney, D. R. (1972). The effect of selected poor item-writing practices on test difficulty, reliability and validity. *Journal of Educational Measurement*, 9, 225-233.
- California Association of Criminalists (May 6, 1994). *Results of the Firearm and Toolmark Examiner Occupational Analysis Survey: an interim report*.
- California Association of Criminalists and Cooperative Personnel Services. (September 27, 1988). *Tasks, Knowledges, Skills and Abilities for Criminalists: A Report*. Sacramento: California.
- Cole, N. (1981). *Bias in testing*. *American Psychologists*, 36, 1067-1077.
- Cooperative Personnel Services, (1997). *The ABC's of Item Writing and Review*, CPS: Sacramento, California.
- Council on Licensure, Enforcement and Regulation. (1993) *Development, Administration, Scoring and Reporting of Credentialing Examinations: Recommendations for Board Members*, First Edition.
- Crocker, L., and Algina, J. (1986) *Introduction to Classical and Modern Test Theory*. New York: CBS College Publishing.
- Ebel, R. L. (1979). *Essentials of Educational Measurement* (3rd Edition). New Jersey: Prentice-Hall.
- Green, K. (1984). "Effects of item characteristics on multiple-choice item difficulty," *Educational and Psychological Measurement*, 44, 551-561.
- Gronlund, N. E. (1968). *Constructing Achievement Tests*. Englewood Cliffs, New Jersey: Prentice-Hall.

- Haladyna, T.M., (1994). *Developing and Validating Multiple-Choice Test Items*. Lawrence Erlbaum Associates: Hillsdale, New Jersey.
- Jensen, A. R. (1980). *Bias in Mental Testing*. New York: The Free Press.
- Linn, R. L. (1989). *Educational Measurement*. New York: Macmillan Publishing Corporation, Third edition.
- Livingston, Samuel A., and Zieky, Michael J. (1982). *Passing Scores: A Manual for Setting Standards of Performance on Educational and Occupational Tests*, Princeton, NJ: Educational Testing Service.
- McMorris, R. F., Brown, J. A., Snyder, G. W., & Pruzek, R. M. (1972). "Effects of violating item construction principles," *Journal of Educational Measurement*, 9, 287-295.
- Millman, Jason, and Greene, Jennifer. (1989). The specification and development of tests of achievement and ability. In R.L. Linn's *Educational Measurement*. New York: Macmillan Publishing Corporation.
- Osterlind, S.J. (1989). *Constructing Test Items*. Kluwer Academic Publishers: Boston.
- Schmeiser, C. B., & Whitney, D. R. (1975). Effect of two selected item writing practices on test difficulty, discrimination, and reliability. *The Journal of Experimental Education*, 43, 3, 30-34.
- Schmitt, K.L. (1998). Issues in Standard Setting. Paper presented at the Society for Industrial and Organizational Psychology Annual Conference, April 26, 1998, Dallas, TX.
- Society for Industrial and Organizational Psychology. (1987). *Principles for the Validation and Use of Personnel Selection Procedures*, third edition, SIOP: College Park, Maryland
- Thorndike, R. L. & Hagen, E. (1969). *Measurement and Evaluation in Psychology and Education*. New York: Wiley.
- U.S. Equal Employment Opportunity Commission, U.S. Civil Service Commission, U.S. Department of Justice, U.S. Department of Labor (1978). Uniform Guidelines on Employee Selection Procedures. *Federal Register*, 43, 38290-38315.
- U.S. Department of Labor. (1995). *United States Employment Service Test Development Manual*, Washington, D.C.
- Weiten, W. (1984). Violation of selected item construction principles in educational measurement
Journal of Experimental Education, 52, 3, 174-178.