SUPPORTING STATEMENT

UTILITY OF TEST PREPARATION GUIDES AND EDUCATION PROGRAMS IN ENHANCING RECRUIT CANDIDATE PERFORMANCE ON THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY (ASVAB)

A. JUSTIFICATION

1. Need for Information Collection

The 2007 National Defense Authorization Act (NDAA), Section 546, requires the Secretary of Defense to conduct a test of the utility of commercially available test preparation guides and education programs designed to assist recruit candidates achieve scores on military recruit qualification testing that better reflect the full potential of those recruit candidates in terms of aptitude and mental category.

This research is needed to better understand the recruit market and to investigate how best to expand the market by providing test preparation materials to prospective recruits.

2. Use of Information

The Department of Defense will use the information to determine the effects of ASVAB test preparation on the prospective recruit market and investigate how best to expand the market by providing test preparation materials to prospective recruits. The results will be reported to the Department of Defense (DoD), Office of Accession Policy, Office of the Under Secretary of Defense for Personnel and Readiness in order to inform DoD wide policy on providing test preparation materials to military service applicants.

The respondents will consist of 145,000 military service applicants who complete the ASVAB when taken at Military Entrance Processing Stations (MEPS) and Military Entrance Testing Sites (METS). The questionnaire responses will be collected in two overlapping data collection efforts:

- One effort involves a 12-month treatment-outcome study of Army National Guard applicants from 5 states up to a count of 20,000. This data collection is designed to meet specific study design parameters specified in the 2007 NDAA. About half of the participating applicants will be assigned to a control treatment condition (TX0), one quarter to treatment manipulation one (TX1) and one quarter to treatment manipulation two (TX2) and half to a control condition.

  If any of the control group members chooses to be retested about half will be assigned to test preparation booklet treatment group (TXR1). See Table 1.

Table 1. Treatment groups and size

Initial test	Retest	Size
TX0-Factsheet		5,000
TX0-Factsheet	TXR1-Book	5,000
TX1-Book		5,000
TX2-Factsheet+Book		5,000
		20,000

- The other effort involves a 3-month, nation-wide questionnaire of all military applicants up to a count of 125,000. The nation-wide questionnaire data collection is designed to meet additional study objectives specified in the 2007 NDAA that are not met by the treatment-outcome study.

The sample size is relatively large because the effects of test preparation can only be assessed if it occurs between the two test sessions and only 15% of applicants take the test two or more times. Of the 145,000 military applicants receiving the questionnaire about 21,750 are expected to re-test. About half of the applicants who re-test are expected to engage in some form of test preparation yielding a treatment group size of 10,000. However, the effects of eight different sources of test preparation materials are going to be statistically evaluated and the rate that these sources are utilized will vary from frequent (e.g., study guide book) to infrequent (e.g., Classroom-based ASVAB preparation course). In addition, the effects of level-of-effort (amount of time that applicants spend in test preparation) are going to be statistically evaluated for various demographic groups and applicants of various aptitude levels. Given the complexity of the analysis plan it is expected that some useful analyses will be based on as few as 188 applicants with sufficient power to detect moderately sized effects (See Section B1, Table 5 and Section B2, Table 6).

The instrument used to collect the information is the ASVAB Preparation Questionnaire which has received a scientific review and approval by our human subjects institutional review board (IRB). The questionnaire will be administered electronically in the MEPS and via paper and pencil in the METS where computers are not available for ASVAB or questionnaire administrations.

The objectives of the study are to determine the following with respect to ASVAB-specific test preparation:

- Does such test preparation help potential military applicants achieve higher scores without sacrificing the reliability and accuracy of the enlistment test?
  To answer this question we plan to compare ASVAB retest scores of applicants who receive test-prep books (Treatment groups) with applicants who do not receive test-prep books (Control group). We also plan to compare correlations of ASVAB scores and training school final scores of treatment and control groups[1].
- Which segments of the recruiting market would benefit from test preparation?
  To answer this question we plan to evaluate the effects of various demographics available from the MEPCOM record on the results of the previous analyses: Home State, Home Zip Code, Citizenship, Gender, Marital Status, Number of Dependents, Race, Ethnicity, Age, Education level, Proficiency in foreign language.

[1] Training school final scores are difficult to collect since they do not reside in one central location. Thus, this part of the analysis plan will be conducted in a separate follow-on study if meaningful test preparation effects are found.

- Which segments of the recruiting market would be more motivated to engage in test preparation if recruiters provided free test preparation materials?
  To answer this question we plan to evaluate the effects of various demographics available from the MEPCOM record on the ASVAB Preparation Questionnaire on the responses to motivational questions on the ASVAB Preparation Questionnaire.
- The degree to which test preparation assistance affects test reliability and accuracy.
  To answer this question we plan to compare correlations of ASVAB scores and training school final scores of treatment and control groups[1].
- Determine the extent to which test preparation assistance allows more accurate testing of skill aptitudes and cognitive capability.
  To answer this question we plan to quantify the magnitude of the correlations of ASVAB scores and training test scores of treatment and control groups.
- What role should be recommended for test preparation assistance in military recruiting?
  To answer this question we plan to provide the policy office with a report on the results of the study.

Structural Equations Modeling (SEM) of the ARNG data will be used to test for construct validity and standard errors of measurement (i.e., reliability). These models will test for differences in the internal structure of ASVAB subtests between pre and post ASVAB scores and between treatment and control groups. Details for the Reliability Model are attached.

Tests of predictive validity will be considered if the results of the treatment-outcome study reveal a significant and substantial effect for the distribution of test-prep books. Although recruiters and other leaders in the recruiting community believe that handing out test-prep books will pay off, evaluations of more intensive test preparation in the private sector suggest otherwise. Powers & Rock (1998)[2] show that the largest effects of major test prep programs (40 hours classroom + 10-20 hours homework) are relatively modest. For example, three or four additional questions correct on the 60-question math portion of the SAT I. Power & Rock (1998) also report that their results are consistent with the results of several previous meta-analyses of the effects of coaching on an earlier version of the SAT.

Reliable and valid training data will be needed to conduct an evaluation of criterion validity. These data are collected in the military training schools for the purpose of administering training but not for research purposes. Assembling and cleaning training data will require months of effort by several researchers. Most importantly, analyses will need to be performed within training schools in order to preserve the reliability and validity of the training measures. Consequently, the sample sizes of students who participate in the study and attend the same training schools may be too small to support analyses of sufficient statistical power.

Although Sec. 546 of the NDAA specifically mentions assessing duty performance for 18 months following active duty, longitudinal criterion validity studies will probably need to wait until the distribution of test-prep books becomes more wide-spread which may not occur if this study does not reveal substantial effects for this practice.

Self-selection is operating in the nation-wide survey. We recognize that minimizing the impact of self-selection in our analyses is vitally important and we will include caveats about self-selection bias

with our results. Our plan is to use Analysis of Covariance (ANCOVA) to control for differences between test-prep users and non-users. A report by Powers & Rock (1998) evaluated seven different

[2] Powers, D. E. & Rock, D. A. (1998). Effects of coaching on SAT I: Reasoning scores. New York: College Entrance Examination Board.

analysis models designed to correct for self-selection bias and suggest that ANCOVA is accurate enough for the purpose of our study. As a methodological check we plan to use some of the additional analysis methods employed by the College Board (i.e., propensity matching, and Heckman modeling) on a selective basis. Every effort will be made to introduce covariates to our models in order to control for differences between those who prepare for the ASVAB and those who do not prepare.

At the conclusion of the study, the Secretary of Defense shall submit to the Committee on Armed Services of the Senate and the Committee on Armed Services of the House of Representatives a report providing the findings of the Secretary with respect to each of the objectives specified above and the Secretary's recommendations.

3. Improved Information Technology

Electronic collection of the questionnaire is employed when the ASVAB is administered electronically, which will occur in approximately 50% of the questionnaire administrations. In this instance the questionnaire will be administered on the same computer. The data will be off-loaded through standard data accumulation procedures used for examinee test records.

Paper and pencil submission of the questionnaire is employed when the ASVAB is administered by paper and pencil because computers are not available in these settings.

4. Efforts to Identify Duplication

There is no information available elsewhere that can be used to evaluate the usefulness of commercially available test preparation guides and education programs designed to assist recruit candidates achieve scores on military recruit qualification testing. Efforts to identify duplicate sources of information included a review of studies, and information found through contacts with industry trade associations, governmental agencies, and academic researchers.

5. Methods Used to Minimize Burden on Small Entities

Small entities are not included in the sample. The questionnaire is only being administered to individual military applicants who complete the ASVAB when taken at MEPS and METS.

6. Consequences of Not Collecting the Information

If this information is not collected, the DoD will be in non-compliance with the 2007 National Defense Authorization Act (Sec. 546). In addition the DoD will be unable to make informed policy decisions regarding test preparation assistance in military recruiting.

7. Special Circumstances

The collection of information in the proposed questionnaire does not incur any of the special circumstances outlined in the narrative of Question 7 of the "General Instructions" of the OMB Form 83-I.

8. Agency 60-Day Federal Register Notice and Consultations Outside the Agency

A 60-day notice was published in the Federal Register on May 2, 2008 (73 FR 24254-24255). No comments were received.

Consultations were conducted with:

Paperwork Reduction Act Submission Review
Cindy L. Allard, Chief, OSD/JS Privacy Office

Human Subjects Protection
Kimberly L. Odam, MS, CCRP, Human Subjects Protection Scientist (AMDEX Corp.)
Andrea Zucker, Exempt Determination Official, For the ODUSD Program Integration
Francine R. Jones, CIP, Northrop Grumman Corporation, Contractor to Force Health Protection and Readiness Programs TRICARE Management Activity Office of the Assistant Secretary
Caroline Miner, HRPP Program Manager for the OUSD(P&R), FHP&RP/TMA/OASD(HA)

ASVAB Test Preparation Study - Research Design Meeting, June 19, 2007, 1230, Room 2B271, The Pentagon:

Dr. Jane Arabian, Assistant Director, Accession Policy
Ms. Janet Held, US Navy Personnel Research Studies & Technology
Ms. Denise Mills, Army
Dr. Lisa Mills, Air Force
Mr. Mike Styka, Marine Corps
Dr. Steve Watson, Navy Selection and Classification Office
Dr. Mike Rumsey, Army
Dr. Len White, US Army Research Institute
Dr. Curt Gilroy, OSD
CMSgt Mike Gasporetto, USAF
Michael Collins, Army
Captain Dan Gilbert, Army
Mr. Jeffrey Spara, Army
Dr. Tom Caretta, Air Force Research Laboratory
Dr. Paul DiTullio, Air Force Reserves
Mr. Brian Labashosky, USAREC
Mr. Don Hill, USMEPCOM
Mr. Bob Philips, CNRC

Human Resources Research Organization
John P. Campbell
Teresa L. Russell
Peter Ramsberger
Shonna D. Waters

9. Payments to Respondents

There will be no compensation for completing the questionnaire or participation in the study.

10. Assurance of Confidentiality

Privacy act data (i.e., SSN and name) collected is covered under a system of records held by Defense Manpower Data Center: DMDC-08, Questionnaire and Census Data Base. This data will be kept private to the extent permitted by law under the Privacy Act.

SSN is needed for matching questionnaire responses with ASVAB test scores. Last name is also included to provide an alternative approach to matching when the SSNs on some records do not match back to the ASVAB score records. SSN provides the only reliable means for making a positive match between questionnaire responses and ASVAB score records.

11. Sensitive Questions

Parental education levels represent a sensitive question on the questionnaire. Collection of this information is also covered under a system of records held by Defense Manpower Data Center: DMDC-08, Questionnaire and Census Data Base.

Parental education level is needed as a proxy for social economic status (SES). SES is correlated with ASVAB scores and possibly ASVAB preparation behaviors. Thus SES may play a key role in the analysis of the relationship between ASVAB scores and preparation behaviors.

Security Measures for Collection and Storage of Data

- Personal identifiers collected via paper and pencil administrations at the METS will be handled per US MEPCOM Records Management procedures for sending questionnaire returns to the MEPS. Shipping questionnaire returns from the MEPS to the operations contractor for scanning and dataset creation will also conform to these procedures. That is:

  Controlled test material will be double-wrapped and mailed using a traceable delivery method. The inner envelope or package (also addressed) will be stamped with the following notation: FOR OFFICIAL USE ONLY CONTROLLED TEST MATERIAL TO BE OPENED ONLY BY APT TEST CONTROL OFFICER. The outer package will not indicate it contains controlled test material or it is FOUO.

- Personal identifiers collected via computer administrations at the MEPS will be retrieved using the same secure computerized methods used to accumulate and retrieve ASVAB test results.
  - Access to the data collection system will be via password-protected computers and local area networks protected by secure firewalls.
  - All data delivered by the contractor will be encrypted and encryption keys delivered via alternate methods.
  - All data held by the contractor will be destroyed 1 month after completion of the study and written acceptance of study materials and reports by the government.
  - Datasets kept by the government will be held on secure password-protected computer systems which are linked to the internet via secure firewalls.

- Analysis datasets will be constructed by matching the contractor datasets with complete ASVAB testing results and miscellaneous demographic data held by the government. SSNs will be removed from the analysis datasets and replaced by a unique random identifier.

12. Estimates of Annual Response Burden and Labor Cost for Hour Burden to the Respondent for Collection of Information.

a. Response Burden:

i. Respondents (frequency 1 time)

Total annual respondents:	123,250
Frequency of Response:	1
Total annual responses:	123,250
Burden per response:	14 minutes
Total burden hours	28,758 hours

ii. Respondents (frequency 2 times)

Total annual respondents:	21,750
Frequency of Response:	2
Total annual responses:	43,500
Burden per response:	14 minutes
Total burden hours	10,150 hours

iii. Respondents (Annual Average)

Total annual respondents:	166,750
Burden per response:	14 minutes
Total burden hours	38,908 hours

b. Explanation of how burden was estimated.

A pre-test of the questionnaire package with seven military applicants was conducted to see how long it would take to complete the questionnaire and to identify wording changes. The questionnaire pamphlet incorporated the Agency Disclosure Notice, Privacy Act Statement, questionnaire introduction, questionnaire instructions, and questionnaire content. The results in Table 2 show that the average time to complete the questionnaire is 13 minutes and 34 seconds (13.57 minutes). Only one response entered the Skip-pattern. The maximum time allowed is 20 minutes.

Table 2. Pre-Test Time to Complete Questionnaire

Subject	Minutes
1	14
2	14
3	10
4	11
5	9
6	20
7	17
Average	13.57

c. Labor Cost to Respondent.

The labor cost to respondent is estimated to be $4.02 [$17.24 per hour (GS-7/5 hourly rate) times 14 min divided by 60 per hour].

Thus, the labor cost for all responses is estimated to be 166,750 times $4.02= $670,335

13. Estimates of Other Cost Burden for the Respondent for Collection of Information.

a. Total Capital and Start-up Cost.

There are no capital/startup costs.

b. Operation and Maintenance Cost.

There are no operation and maintenance costs.

14. Estimates of Cost to the Federal Government.

The total cost to the Federal Government for the information collection will be approximately $729,350.

To conduct the questionnaire data collection portion of the study (i.e., a. print and distribute the questionnaires, b. administer the questionnaires, and c. analyze the data) is estimated to cost $373,695. Clerical and other support services and costs are included in this estimate.

To conduct the treatment outcome experiment (i.e., a. write the contract, b. manage the contract, and c. purchase and distribute test preparation books) is estimated to be $355,655. Clerical and other support services and costs are included in this estimate.

15. Changes in Burden.

Increase in burden is due to a new information collection.

16. Publication Plans/Time Schedule.

The results of this study will be used to prepare a public report to the Congress. The collection of the information will begin within two months of the completion of the OMB review process. These two

months will be used for the printing & distribution of the paper version of the questionnaire and the distribution and testing of the electronic version of the questionnaire. The projected duration of the information collection is approximately twelve months. The estimated date for the completion of the report is early 2009.

17. Approval Not to Display Expiration Date.

Approval Not to Display Expiration Date is not needed.

18. Exceptions to the Certification Statement.

There are no exceptions to the certification statement.

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Description of Sampling Methodology

The potential respondent universe consists of all military applicants who complete the ASVAB when taken at Military Entrance Processing Stations (MEPS) and Military Entrance Testing Sites (METS). The sampling plan is a convenience census of a subset of the universe and is considered sufficient for the purposes of this study. ASVAB scores and demographic data on the non-covered population which is available on the MEPCOM record will be available to evaluate the representativeness of the sample.

The nation-wide questionnaire will be conducted for three months to yield a sample of approximately 125,000 applicants. The sample size of 125,000 was designed to meet the statistical requirements of in-depth analyses. This sampling plan has the potential of incurring non-negligible coverage error; however, this will not thwart the purpose of the study.

Table 3. Universe and Sample Counts for Nation-wide Questionnaire

Universe	Census
12 Months	3 Month
500,000	125,000

The Army National Guard (ARNG) treatment outcome study will run for twelve months in five states which volunteered to participate in the study to a sample of about 20,000 applicants. The sample size of 20,000 was arbitrarily set according to the requirements in the NDAA of 2007. The NDAA requires that data are collected from at least 2,000 treatment group and 2,000 control group subjects and we are assuming that at least 20% of the military applicants recruited into the study will actually complete the ASVAB and the ASVAB Preparation Questionnaire. The assignment of examinees to treatment and control groups will be pseudo randomized based on the last four and last three digits of the SSN. The last four digits will be used to assign subjects to the initial-test treatment groups and the last three digits will be used to assign subjects to the retest treatment groups.

Table 4. Universe and Sample Counts for ARNG Treatment Outcome Study

Universe	Census
All States	5 State
67,084	20,000

In ASVAB testing sessions nearly 100% of examinees complete the ASVAB. A similar response rate for the ASVAB Preparation Behavior Questionnaire is expected (No voluntary questionnaires have been administered with the ASVAB, therefore applicable response rates are not available).

2. Describe the procedures for the collection of information including:

The sample size requirement of 125,000 military applicants for the nation-wide questionnaire was based on the outcome of a statistical power analysis. Table 5 displays the expected sample sizes from different hypothetical test prep behaviors for a three month data collection effort. These for test prep behaviors are assumed to be exhibited by different fractions of population, ranging from very common (1 out of 5 applicants) to very infrequent (1 out of every 100 applicants). These expected sample sizes are based on the following assumptions:

- 500,000 applicants test annually
- 15% of applicants take the test two or more times
- Among those who take the ASVAB more than once and engage in test prep, about half will engage in test prep between the most recent and previous tests. (The other half of applicants will have engaged in test prep prior to the previous or first ASVAB taken.)
- The data collection will last three months.
- Control Group Sample Size = 500,000 x 0.15 x 3/12 x 1/2 = 9,375
  (Assumes that 15% of the sample takes the test more than once, data collection lasts for 3 months, and half the population does not engage in any form of test prep.)
- Treatment Group Sample Size = 500,000 x 0.15 x 3/12 x 1/2 x F
  (Assumes that 15% of the sample retests, data collection last for 3 months, among those taking the test twice and engaging in test prep–half engage just prior to the most recent test, and *F* is the proportion of the population engaging in the target test-prep behavior.)

As indicated in Table 5, given the above assumptions we expect a Control Group sample size of 9,375. For relatively frequent test prep activities (say those involving 1 out of every 5 examinees–like familiarizing oneself with the test lengths and time limits) we would expect a Treatment Group of about 1,875. Conversely, for relatively infrequent test prep activities (say those involving 1 out of every 50 examinees–like participating in a several week instructor led course) we would expect a Treatment Group of about 188.

Table 5. Expected Sample Sizes based on fraction engaging in test pre behavior.

Population Fraction (F) (*F*)	Sample Size	
	Treatment	Contro
1/5	1,875	9,375
1/10	938	9,375
1/15	625	9,375
1/20	469	9,375
1/25	375	9,375
1/40	234	9,375
1/50	188	9,375

The power to detect mean differences in ASVAB performance (between the Control and Treatment groups) is displayed in Table 6. This table displays power for two different effect sizes (0.1 standard deviation units and 0.2 standard units), assumes a one tail *t* test, a significance level $\alpha = 0.05$, and assumes equal group variances. The expected samples sizes listed in Table 5 will provide high power to detect small (but meaningful) subtest effects (column 2, 0.1 SD Units) for relatively frequent test preparation activities (those engaged in by 10% or more of the population). The planned sample sizes will also provide high power to detect a moderate subtest effect size (column 3, 0.2 SD Units) for less frequent test prep behaviors (those engaged in by 2% or more of the test-taker population). The same results are true for the Armed Forces Qualification Test (AFQT) effects (Table 6, columns 4 and 5).

Table 6. Power to detect increase in mean test scores due to test

Population fraction	Subtest Effect Size		AFQT Effect Size	
	0.1 SD Units	0.2 SD Units	0.1 SD Units	0.2 SD Units
1/5	0.99	1.00	0.99	1.00
1/10	0.90	1.00	0.93	1.00
1/15	0.78	1.00	0.82	1.00
1/20	0.68	1.00	0.72	1.00
1/25	0.60	0.98	0.64	0.99
1/40	0.45	0.92	0.48	0.94
1/50	0.39	0.86	0.42	0.89

3. Describe methods to maximize response rates and to deal with issues of non-response.

High response rates are expected due to the setting of questionnaire administration. The questionnaire instructions and content were carefully constructed in order to maximize completion rates.

4. Describe any tests of procedures or methods to be undertaken.

A pretest was conducted on eight staff members prior to a scientific review of the questionnaire in preparation for the protection of human subjects review in order to refine the survey and estimate expected response times. Seven military applicants were also surveyed in order to further refine the questionnaire. Additional test administrations to military applicants may be undertaken but they will not cause the total number of administrations to exceed the burden stipulated in this submission.

5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Daniel Segall, Ph.D.
Chief of Personnel Testing Division
Defense Manpower Data Center (DMDC)
400 Gigling Rd
Seaside CA 93955
(831)583-2400
dan.segall@osd.pentagon.mil

Richard Riemer, Ph.D.
Personnel Testing Division Analyst
Defense Manpower Data Center (DMDC)
400 Gigling Rd
Seaside CA 93955
(831)583-2400
richard.riemer@osd.pentagon.mil

Jeff Barnes
Human Resources Research Organization
Alexandria, Virginia
703-706-5691
JBarnes@HumRRO.org

Richard Branch
Command Technical Advisor/Chief, Test Score Control Office
HQ USMEPCOM/MIT-TSCO
richard.branch@mepcom.army.mil

Lloyd Bowers
SFC, ARNG
COR
1411 Jefferson Davis Highway
Arlington, VA 22202
(703) 601-2993
lloyd.j.bowers@us.army.mil