OMB questions and DoD answers about the Utility of Test Preparation Guides and Education Programs in Enhancing Recruit Candidate performance on the ASVAB.

Questions about the Utility of Test Preparation Guides and Education Programs in Enhancing Recruit Candidate performance on the ASVAB

In A.10 of the supporting statement, there is no mention of any statutory authority for a pledge of confidentiality. However, the questionnaire states that responses will be "confidential." Please provide the citation and excerpt of the statutory authority for this promise or modify the language in the questionnaire to state that the information will be kept "private."

> We will change the indications of confidentiality to private.

In A.2 of the supporting statement, DoD provided a list of objectives for the study; however, it was not clear from the materials submitted how the proposed study will be able to answer each of the research questions underlying these objectives. For each objective, please provide more information the design of that component of the study and the analysis plans showing how the data gathered in this study will be used to answer each of these research questions.

- Does such test preparation help potential military applicants achieve higher scores without sacrificing the reliability and accuracy of the enlistment test?
  >Compare ASVAB retest scores of applicants who receive test-prep books (Treatment groups) with applicants who do not receive test-prep books (Control group)
  >Compare correlations of ASVAB scores and training test scores of treatment and control groups.
- Which segments of the recruiting market would benefit from test preparation?
  >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.
- Which segments of the recruiting market would be more motivated to engage in test preparation if recruiters provided free test preparation materials?
  >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.
  >Compare correlations of ASVAB scores and training test scores of treatment and control groups.
- Determine the extent to which test preparation assistance allows more accurate testing of skill aptitudes and cognitive capability.
  >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?
  >Provide the policy office with a report on the results of the study.

It is not clear if the ARNG treatment outcome study is randomly assigning potential recruits to treatment and control groups or how this is being done

> 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.

there is no discussion of what measures of reliability and validity are being employed or how the study will examine the effects of test preparation on these measures.

>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.

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) 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.

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

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. Criterion validity analyses 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.

Similarly, it appears from the description that potential recruits are self-selecting different preparation methods, so it is not clear what conclusions can be drawn from this analysis about the efficacy of such methods.

>Random assignment is used in the ARNG treatment-outcome study which will show if applicants achieve higher scores (and/or larger retest gains) when they receive ASVAB preparation books.

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 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.

In addition, Sec. 546 of the NDAA specifically mentions assessing duty performance for 18 months following active duty, but I saw no mention of longitudinal follow-up anywhere in the materials submitted.

>Training performance will be used if significant test-prep effects are found and sample sizes for homogenous training groups are large enough to detect meaningful effects. Training performance measures are available and generally have higher reliability and validity than available on-the-job

performance measures. Creating research quality on-the-job performance measures would be a time intensive long-term endeavor and at this time no funding has been made available to cover the cost.

In B.1, the sampling plan is described as a "convenience census," but in B.2.c., it was noted that you will take into account any complex sampling or post-stratification through Taylor series estimation. Please describe the complex sampling design and the weighting adjustments (including post-stratification) that will be made for estimates from this study.

>A cluster sample design of months for the nation-wide survey and states for the ARNG study was developed, but given the added difficulty of implementing those designs and the availability of demographic variables for post-stratification it was decided to pursue a less complex sampling method. As noted above we will have the following demographic variables available for minimizing coverage errors: Home State, Home Zip Code, Citizenship, Gender, Marital Status, Number of Dependents, Race, Ethnicity, Age, Education level, Proficiency in foreign language.

What is the expected response rate for the questionnaire? Have other voluntary questionnaires been administered with the ASVAB before? If so, what response rates were achieved with those?

>No voluntary questionnaires have been administered with the ASVAB therefore applicable response rates are not available.

The questionnaire should be pretested with actual respondents to ensure appropriate understanding of question wording, and to assess flow and usability of the instrument.

>We had scheduled a pretest of nine examinees on June 25th. But MEPCOM has not given us permission at this time. This type of request is not unusual, but there has been a delay as MEPCOM reviews the instructions to the Test Administrators.

Questionnaire items:
It is not clear what 6a-n are referring to. Is this information in instructions or a brochure? It is also not clear why they are lumped together with items o to x, which would seem to be the key items you want to examine in this study.

>We have revised the wording on this question.

Item 10 seems unclear. Don't you want to more directly ask how much time they spent studying these different content areas with test preparation materials or study aids?

>We have revised the wording on this question.

How would respondents know how they compare to other military applicants (question 18)?

>This question is intended to measure respondent perceptions as they may relate to their motivation to do well on the ASVAB. These types of questions have been asked in similar studies.