NLSY

ATTACHMENT 106: PROFILES

Profiles of American Youth

General & Technical Information on the 1980 Administration of the Armed Services Vocational Aptitude Battery (ASVAB) to Respondents of the National Longitudinal Survey of Youth

ATTACHMENT 106: PROFILES
Background and General Information

Profiles is the term used to describe the study called the Profile of American Youth that was conducted with the NLS youth sample during July through October of 1980. In that study, the Armed Services Vocational Aptitude Battery (ASVAB) was administered to a total of 11,914 NLS respondents, representing a completion rate of approximately 94%. The study was sponsored by the Department of Defense and the Military Services in cooperation with the Department of Labor.

The purpose of the Profile of American Youth was to obtain data on the vocational aptitudes of current youth and to establish current national norms for the ASVAB. Previously, military recruits had been compared statistically to adult males who were extensively tested during World War II. The Department of Defense and the Congress, after questioning the appropriateness of using the World War II reference population as the primary basis for interpreting the enlistment test scores of contemporary recruits, decided in 1979 to conduct this new study.

The NLS respondents were selected because they comprised an already existing nationally representative sample of young people ages 16 to 23. The test was administered by NORC representatives according to strict guidelines conforming to standard ASVAB procedures. Groups of 5 to 10 persons were tested at more than 400 test sites that included hotels, community centers, and libraries throughout the United States and abroad.

For their participation, NLS respondents were paid an honorarium of $50 for completing the test to compensate them for their time and travel expenses. Additionally, the respondents were sent their test results along with information on how to interpret the results and how to use the ASVAB scores in helping to make career choices.

The ASVAB consists of a battery of 10 tests that measure knowledge and skill in the following areas: (1) general science, (2) arithmetic reasoning,

Attachment 106: Profiles continued

(3) word knowledge, (4) paragraph comprehension, (5) numerical operations, (6) coding speed, (7) auto and shop information, (8) mathematics knowledge, (9) mechanical comprehension, and (10) electronics information. The NLS tapes contain 33 Profiles variables, R(6147.) through R(6179.), which include individual number correct scores, scale scores, and standard errors for each of the 10 sections along with sampling weight, test disposition, and high school graduate status at Profiles testing. (Please note that the scale scores include both positive and negative values and that non-interviews have a value (-5) within the range of valid scores.)

Military Services uses ASVAB scores to determine eligibility and assignment qualifications for specific military jobs for new enlistees. Also, Military Services sums the scores of four of the ASVAB subtest scores--word knowledge, arithmetic reasoning, pargraph comprehension and 1/2 of the numeric operations--to create the Armed Forces Qualification Test (AFQT). The AFQT is a general measure of trainability and a primary criteria of enlistment eligibility for the Armed Forces.

Included in this attachment are the following:

(1) Technical notes on the ASVAB scale scores
(2) An annotated bibliography of Department of Defense publications related to the Profile of American youth
(3) The brochure that NORC sent to NLS respondents providing information about the Profiles study and their participation in it
(4) An example of the test score report and related brochures that the participating respondents received after the tests were graded

THE PROFILE OF AMERICAN YOUTH

ASVAB SCALE SCORES

TECHNICAL NOTE

The public release tape of data from the National Longitudinal Survey of Labor Force Behavior, Youth Survey, includes information from the 1980 administration of the Armed Services Vocational Aptitude Battery (ASVAB) to the NLS sample. The scale scores and associated standard errors of measurement are the by-product of psychometric analyses of these data, obtained through the item response methods described in Data Quality Analysis of the Armed Services Vocational Aptitude Battery (R.Darrell Bock and Robert J. Mislevy, Chicago: National Opinion Research Center, 1981). Relevant excerpts of that report are attached.

Inasmuch as neither item content nor item-response data has been released at this time, final estimates have been re-standardized within each ASVAB subtest to weighted population means of zero and standard deviations of one. No scale scores are reported for the 36 subjects who were administered the ASVAB under non-standard conditions.

Full documentation of the sampling methods and fieldwork methods for the administration of the ASVAB test are available in two reports from NORC The Profile of American Youth: Field Report (Harold A. McWilliams, Chicago: NORC, 1980) and The Profile of American Youth: Technical Sampling Report (Martin R. Frankel, Chicago: NORC, 1981).

Excerpted from: DATA QUALITY ANALYSIS OF THE ARMED SERVICES VOCATIONAL APTITUDE BATTERY by R. Darrell Bock and Robert J. Mislevy, National Opinion Research Center, 1981

PART I

DATA-QUALITY ANALYSIS WITH ITEM RESPONSE CURVE MODELS

In order to discover and describe the manner in which each of the ten ASVAB subtests operates to measure subject abilities, each subtest has been fit separately to an item response curve (IRC) psychometric model. This part of our report briefly describes the IRC models that are used: a 3-parameter logistic model for the power (unspeeded) tests and a Poisson model for the speeded subtests.

ITEM RESPONSE CURVE MODELS AND DATA ANALYSIS

The use of item response curve (IRC) psychometric models has long been supported and developed by the armed services, beginning in the 1960's with the work of Allan Birnbaum and continuing in the present with the work of Fumiko Samejima, James McBride, and others. The principal advantages of these models in selection and classification stem from the fact that subject abilities are estimated conditionally on the items the subject has been presented. Tailored testing and linking of test forms, difficult problems for the methods of classical test theory, become quite tractible.

Our choice of IRC models for the analysis of the ASVAB has been motivated from a different point of view. In parameterizing the expected patterns of subject/item interactions, IRC models are excellent vehicles for data analysis of mental test data. To the extent that an IRC explains the data, it expresses the operations of the test and of the items as measuring devices. The amount of information provided about subjects at various levels of ability, by individual items and by the test as a whole, are easily obtained. Departures from the model highlight disturbances and anomalies in the data, such as ambiguities in test items or random guessing behavior of subjects.

Eight of the ten ASVAB subtests are power tests, in which performance depends mainly on subjects' knowledge or reasoning abilities rather than time limitations. To these tests, a Birnbaum 3-parameter logistic IRC (Birnbaum, 1968) was fit with the BILOG computer program (Bock & Mislevy, 1980), using the fixed-effects algorithm outlined in Appendix A.

The remaining two subtests are speeded tests, in which performance depends mainly on subjects' speed and accuracy at simple tasks in a limited amount of time. To both of these tests, a Poisson IRC for speeded tests (Rasch, 1960) was fit.

The following sections review these models in turn.

A 3-PARAMETER LOGISTIC MODEL FOR POWER SUBTESTS

An unspeeded (power) mental test demands knowledge or reasoning ability from a subject. Although time limits are set and observed, it is assumed that subjects would answer few additional items correctly under more generous time limits. Eight of the ten ASVAB subtests are unspeeded tests: (1) General Science, (2) Arithmetic Reasoning, (3) Word Knowledge, (4) Paragraph Comprehension, (5) Auto & Shop Information, (6) Mathematics Knowledge, (7) Mechanical Comprehension, and (8) Electronics Information.

The model used to analyze the eight power subtests of the ASVAB is based on Birnbaum's 3-parameter logistic item response curve (IRC) model. The IRC provides a statistical model for the probability of a given subject responding correctly to a given test item. Included in this model is the subject's scale score and parameters that characterize the difficulty and reliability of the item. When the model is fitted to the data, it is capable of accounting for the facts that--

(1) Some subjects perform better than others on the items in the subtest.

(2) Some items in the subtest are easier than others.

(3) Some items measure the underlying ability more precisely than others.

(4) Because the test items are multiple choice, subjects can

occasionally answer any item correctly by guessing.

The scale of ability along which persons are measured is defined explicitly by the locations of the items:

The ability, or scale score, of Subject i (θ_i) is his location along the scale.

The location of Item j on the scale is called its threshold (b_j). Items' thresholds indicate their relative probabilities of being answered correctly by a person drawn at random from the target population. A subject located at the threshold of Item j would have a 50-50 chance of answering Item j correctly.

The dispersion parameter of Item j (s_j) is inversely related to the reliability with which item j measures ability.

Finally, the lower asymptote of Item j (c_j) is the probability of a correct response from even the subjects of lowest ability. Lower asymptotes may be useful during the estimation of item parameters, freeing threshold and dispersion estimates from the effects of random guessing. Because guessing behavior differs from one subject to another, however, final estimates of subject ability are instead based on a a robust procedure that does not use the lower asymptotes (see Appendix B).

The exact value of the probability of a correct response to Item j from Subject i is given by the following function of θ_i, b_j, s_j, and c_j:

$$P_{ij} = c_j + (1-c_j) \exp(Z_{ij}) / [1.0 + \exp(Z_{ij})],$$

where

$$z_{ij} = (\theta_i - b_j)/s_j$$

and exp(x) denotes the raising of the base of the natural logarithms to the xth power.

The origin and scale of the ability variable may be chosen arbitrarily. In our analyses, the scale has been set so that the mean of subject abilities in the youth population is zero and the variance is one, after correction for measurement error variation.

The amount of information that Item j provides about subjects at various levels of ability is given by its information curve, which for the logistic item response model is the derivative of the IRC (i.e., the slope of the curve at each point on the scale). The information provided by Item j for ability level θ is given by

$$I_j(\theta) = \{[1+\exp(z_{\theta j})]^{-1}/(s_j)^2\}\{1-c_j\}\{[1+\exp(z_{\theta j} - \ln(c_j)]^{-1}\}$$

where

$$z_{\theta j} = (\theta - b_j)/s_j.$$

It may be inferred that an item provides most information about subjects whose abilities lie in the neighborhood of its threshold. It may also be inferred that, at their most informative points, items with large dispersions provide less information than items with small dispersions.

The total amount of information provided by a collection of items is given by the sum of their individual item information curves, i.e., the test information curve. When the method of maximum likelihood is used to estimate subjects' abilities from their responses to items with known parameters, the standard error

of estimation will be the square root of the reciprocal of the total test information at the estimated ability. An examination of a test's information curve, then, may be used to examine the levels of measurement precision that are attained at various points along the ability scale. Such analyses will be performed for the power tests of the ASVAB.

We note in passing that simpler item-response models with constant item dispersions and/or lower asymptotes of zero have also been proposed (see Lord & Novick, 1968, and Andersen, 1980). These models offer considerable conceptual and technical advantages in applied settings, and are worthwhile goals during test construction. Preliminary analyses of the data from the ASVAB subtests indicated that these simpler models fit the data poorly, however. The 3-parameter model has been adopted as better suited to the task at hand, namely, data analysis of responses to existing tests.

A POISSON MODEL FOR SPEEDED SUBTESTS

The items of the speeded subtests of the ASVAB, Numerical Operations and Coding Speed, require little knowledge or reasoning ability. If time were not restricted so severely, almost every subject would answer almost every item correctly. Performance in these subtests, then, places a premium on subjects' speed and accuracy.

The model used to analyze these subtests is based on Rasch's (1960) model for speeded tests. Two assumptions are necessary. First, item content within a test is considered to be homogeneous --an assumption fairly well satisfied for Numerical Operations and almost perfectly satisfied for Coding Speed. Second, a subtest is treated as if it were infinitely long--an assumption also well-satisfied in both tests, as time restrictions are sufficiently strict to prevent all but a few subjects from reaching the end of either subtest.

Briefly, the justification of the model is as follows: It is supposed that (1) the probability of Subject i responding correctly to an item during any small time interval Δt depends only on the length of the interval, and that (2) as Δt approaches zero the possible outcomes are essentially either <u>one</u> correct response, with probability P_i, or <u>zero</u> correct responses, with probability $(1-P_i)$. Then the probability that Subject i will correctly answer R_i items over the course of N_i time intervals is approximated by the following Poisson model:

$$\text{Prob}(R_i|N_i) = \frac{(N_i P_i)^{R_i}}{R_i!} \exp(-N_i P_i),$$

where $(N_i P_i)$ may be interpreted as the expected number of correct responses.

Rasch's original model expressed P_i as the product of a term for subject ability and a term for test difficulty, which is essential for the comparison or linking of multiple tests.

Because our attention is focused on only one subtest at a time, this separation is not necessary. Furthermore, because every subject is allotted the same amount of time, N_i may be absorbed with P_i into a single parameter θ_i, the ability of Subject i with respect to the given subtest and standard time limitations.

Under the assumptions outlined above, the maximum likelihood estimate of θ_i is simply R_i and its standard error of estimation is the square root of R_i. It is customary to analyze the logs of number-correct scores of speeded tests, and for this reason we have approximated θ_i by $(R_i + 0.5)$. Subsequent analyses of score distributions revealed that the number-correct scores have approximately normal distributions in the youth population, as do the ability estimates from the power tests, while the logs of the number-right scores do not. The number-right metric has been retained, then, so that score units may be more comparable for the power subtests and the speeded subtests. Like the power tests, the speeded tests have been standardized so that the mean of the population is zero and the variance is one, after correction for measurement error variance.

A facsimile of a test information curve is obtained by plotting the squared reciprocals of standard errors against ability estimates. While the form of the model dictates decreasing information with increasing ability, it is useful to examine the test information curve in relation to the population distibution of ability. Moreover, it will be possible to derive statements about local and overall test reliabilities.

APPENDIX A

CALIBRATION OF POWER SUBTESTS

The items in each ASVAB power subtest were calibrated with the BILOG computer program (Bock & Mislevy, 1981) using an adaptation of the fixed-effects solution introduced in Bock (1976). Item parameter estimates were based on the responses of a 10-percent random sample (1200 subjects) of the NLS data, excluding those subjects not tested under standard administration procedures. This appendix outlines the calibration algorithm.

According to the assumptions of item response curve theory, item parameters are invariant across subjects and could be estimated from any sample, regardless of its ability or demographic features. This assumption is never satisfied perfectly in practice, however, so precautions were taken to guard against biases in item parameter estimates caused by the oversampling of disadvantaged subjects. Rather than weight each subject in the calibration sample equally, we have weighted them in proportion to their NLS sampling weights. The weights have been rescaled to make the sum of subject weights add to 1200, the actual number of subjects. The data from a given subject will be weighted inversely to his probability of being selected. In the following discussion, we use the following terms for a subject's item attempts and correct responses:

N_{ij} = the weighted number of attempts by Subject i to Item j

= W_{ij}, and

R_{ij} = the weighted number of correct responses by Subject i to Item j

= W_{ij} if the response is correct and 0 otherwise.

As a consequence of this weighting, the item-fit and test-fit Chi-square statistics will not be strictly correct. It may be appropriate to adjust them in accordance with a design effect, probabably a value around two like the design effects for many other variables in the NLS survey. The resulting item parameter estimates will, however, correspond more closely to those that would be obtained from the responses of a true, simple random sample of the youth population.

CALIBRATION ALGORITHM

Step 1

An initial estimate of the ability of each subject in the calibration sample is obtained as the logit of correct response:

$$\hat{\theta}_i = \ln(\Sigma R_{ij}/(\Sigma N_{ij}-\Sigma R_{ij}).$$

If all the responses of Subject i are correct or all are incorrect, W_{ij} is added to the number of attempts and $W_{ij}/2$ to the number correct.

Step 2

Based on the provisional ability estimates, the calibration sample is partitioned into ten intervals as follows: The lowest and highest scoring five-percent of subjects are assigned to the

lowest and highest intervals, then the attainment scale between these extremes is divided into ten intervals of equal length.

It is assumed that the abilities of subjects in each interval are sufficiently similar to be approximated by a single interval ability. Numbers-tried and numbers-correct of Interval ℓ are defined by

$N_{\ell j}$ = the weighted number of attempts to Item j by subjects in Interval ℓ, and

$R_{\ell j}$ = the weighted number of correct responses to Item j by subjects in Interval ℓ.

Step 3

The probability of a correct response to Item j from a subject in Interval ℓ is assumed to be given by the logistic ogive:

$$P_{\ell j} = .10 + .90 \ \Psi \ (z_{\ell j})$$
$$= .10 + .90 \ \exp(z_{\ell j})/(1 + \exp(z_{\ell j})),$$

where

$$z_{\ell j} = (\theta_\ell - b_j)/s_j$$
$$= (1.0/s_j)\ \theta_\ell - (b_j/s_j)$$
$$= a_j\ \theta_\ell - c_j.$$

The item parameters b_j and s_j are the item threshold and dispersion discussed above. The re-expression in terms of the item parameters a_j and c_j, the item slope and intercept, simplifies computation.

Assuming the local independence of responses to test items, the probability of observing $R_{\ell j}$ correct responses to Item j from the subjects in Interval ℓ is given by

$$P_{\ell j} = \text{Prob}(R_{\ell j} \mid N_{\ell j}, \theta_\ell)$$

$$= \frac{N_{\ell j}!}{R_{\ell j}!(N_{\ell j}-R_{\ell j})!} P_{\ell j}^{R_{\ell j}} (1-P_{\ell j})^{N_{\ell j}-R_{\ell j}}$$

and the probability of the entire calibration data matrix becomes

$$P = \prod_{1}^{10} \prod_{j}^{n} P_{\ell j}$$

Estimates of the a_j's, c_j's, and θ_ℓ's are chosen to maximize this probability. The log likelihood function is

$$L = \prod_{1}^{10} \prod_{j}^{n} C + R_{\ell j} \ln(P_{\ell j}) + (N_{\ell j}-R_{\ell j}) \ln(P_{\ell j}),$$

where C does not depend on the parameters. The likelihood equations for $\ell = 1, 2, \ldots, 10$ and $j = 1, 2, \ldots, n$ are

$$c_j: \sum_{1} (R_{\ell j} - N_{\ell j} P_{\ell j}) = 0$$

$$a_j: \sum_{1} (R_{\ell j} - N_{\ell j} P_{\ell j}) \theta_\ell = 0$$

$$\theta_\ell: \sum_{j} (R_{\ell j} - N_{\ell j} P_{\ell j}) a_j = 0$$

In order to fix the size and origin of the provisional scale units, the highest and lowest intervals are assigned scores of plus one and minus one respectively. BILOG solves the reduced equations by means of Newton-Raphson iteration.

Step 4

From the provisional item parameters estimated in the preceding step, each subject's scale score is estimated. The appropriate likelihood equation, under the assumption of local independence, is given by

$$\theta_i : \sum (R_{ij} - N_{ij} P_{ij}) a_j = 0,$$

where

$$P_{ij} = .10 + .90 \exp(Z_{ij}) / (1 + \exp(Z_{ij}))$$

with

$$Z_{ij} = a_j \theta_i - c_j.$$

This equation has no solution if all of the responses of Subject i are correct or all are incorrect. In the former case, W_{ij} is added to the number of attempts to the item with the highest threshold and $W_{ij}/2$ is added to his number correct; in the latter case, the same procedure is applied to the item with the lowest threshold.

Step 5

Step 2 is repeated with the improved subject score estimates.

Step 6

Step 3 is repeated with the improved interval boundaries.

Step 7

Standard errors of estimation for the item parameters are obtained in the final Newton iteration of Step 6, as the square roots of the negative reciprocals of the second derivatives of the log likelihood at the final solution.

Item fit is indicated by a Pearsonian chi-square over intervals:

$$X_j^2 = \Sigma \ (R_{\ell j} - N_{\ell j} P_{\ell j})^2 / [N_{\ell j} P_{\ell j} (1-P_{\ell j})].$$

If the expected number of either correct or incorrect responses to Item j in Interval ℓ is less than 5, the Interval ℓ is collapsed into an adjacent interval for the purpose of the item-fit index. The number of degrees of freedom associated with the value is two less than the number of intervals after collapsing.

Overall test fit is indicated by the sum of the item-fit

chi-squares, with degrees of freedom similarly summed but reduced by 8 to account for the estimation of interval scores.

As noted above, it may be appropriate to divide the resulting item and test fit Chi-squares by two to account for the stratified sampling design.

APPENDIX B

BIWEIGHT ESTIMATES OF LATENT ABILITY

Maximum likelihood estimates of subjects' abilities in item response curve models are overly senstive to disturbances that are common in educational measurement, such as guessing and carelessness. The biweight solution described in this appendix, introduced and illustrated in Mislevy and Bock (1980), is highly resistent to these disturbances. It effectively discounts suspicious responses, and agrees with the maximum likelihood estimator when none are present.

We first review the form of maximum likelihood estimation of latent ability. Our final estimates of subject abilities in the power subtests of the ASVAB employ two variations on this basic theme, namely, the use of a prior distribution and biweighting.

MAXIMUM LIKELIHOOD ESTIMATES OF LATENT ABILITY

Suppose that the item parameter of n test items are known. Let b_j be the threshold of Item j and s_j be the dispersion. We observe the responses of Subject i to these items. Let X_{ij} be one if the response to Item j is correct and zero if it is not.

Under the assumptions of the 2-parameter logistic item response curve model, the probability that Subject i will respond correctly to Item j is given by

$$P_{ij} = \exp(Z_{ij}) / (1 + \exp(Z_{ij})), \qquad (B.1)$$

where

$$Z_{ij} = (\theta_i - b_j)/s_j$$

and θ_i denotes the ability of Subject i. Assuming the responses of Subject i are independent, given θ_i, the probability of his vector of responses is given by the product of expressions like (B.1) over all the items:

$$P_i = \prod_{j=1}^{n} P_{ij}^{X_{ij}} (1-P_{ij})^{1-X_{ij}}. \qquad (B.2)$$

If the item parameters are known but the ability is not, then (B.2) is the likelihood function of θ_i given the vector of responses. The maximum likelihood estimate of the ability, $\hat{\theta}_i$, is the value which maximizes (B.2) with respect to the observed responses.

In practice the log of the likelihood is maximized. The log likelihood is given by

$$\ell n\ L = \sum_{j=1}^{n} C + X_{ij}\, \ell n(P_{ij}) + (1-X_{ij})\, \ell n(1-P_{ij}).$$

where C does not depend on the parameters. The first derivative of the log likelihood function is given by

$$\frac{d\ell nL}{d\theta_i} = \sum_{j=1}^{n} (X_{ij}-P_{ij})/s_j. \qquad (B.3)$$

and its second derivative, by

$$\frac{d\ \ell nL}{d\theta_i} = -\sum_{j=1}^{n} P_{ij}\ (1-P_{ij})/\ s_j^{\ 2}.$$

As long as not all of the responses are correct and not all are incorrect, the is a unique and finite value for which the

first derivative is zero. Since the second derivative is always negative, the zero is a maximum of the log likelihood. A large-sample standard error for the estimate may be obtained as the negative reciprocal of the square root of the second derivative of ℓn L, evaluated at the maximum.

BAYES MODAL ESTIMATES

As noted in Part I of this report, the scale of the item parameters was fixed by requiring the mean of the youth population to be zero and the true-score variation to be one. Under the assumption that the underlying distribution is normal, it is possible to use Bayes Theorem to obtain estimates of subject abilities with lower mean-squared errors than maximum likelihood estimates.

Under this scheme, the prior density of θ_i is Normal (0,1). The posterior density, F, is proportional to the prior density times the likelihood (B.2):

$$F = \{ \prod_{j=1}^{n} P_{ij}^{X_{ij}} (1-P_{ij})^{1-X_{ij}} \} \times \{ \frac{-1}{2\Pi} \exp(-\theta_i^2/2) \}$$

where P_{ij} is as defined in (B.1). The value of θ_i that minimizes this expression is the Bayes modal estimate of $\tilde{\theta}_i$, the highest value with the highest posterior density. The log of F and its first and second derivatives are nearly the same as those for L, except for additional terms:

$$\ell n\ F = \sum_{j=1}^{n} K + X_{ij}\, \ell n(P_{ij}) + (1-X_{ij})\, \ell n(1-P_{ij}) - \theta_i^2/2.$$

$$\frac{d\ell nF}{d\theta} = \theta_i + \sum_{j=1}^{n} (X_{ij}-P_{ij})/s_j.$$

$$\frac{d^2\ell nF}{d\theta^2} = -1 - \sum_{j=1}^{n} P_{ij}\,(1-P_{ij})/\,s_j^{\,2}.$$

It is typical to take as an indication of the precision of estimation the negative reciprocal of the square root of the second derivative at the maximum; i.e., the curvature of the posterior distribution at its highest point.

BIWEIGHT ESTIMATES

In theory, a subject's responses to items with thresholds far above or far below his level of ability provide little information about his ability. In practice, they may provide misinformation. An incorrect response to an easy item from a subject who otherwise appears quite able is probably a careless error; a correct response to a hard item from a subject who otherwise appears unable is probably a lucky guess. Inasmuch as a subject's responses to items far from his apparent ability contain least information and most potential for misinformation, it would be desirable to weight a subject's responses accordingly. The biweight estimator described in this section does just that.

Based on the principal of Tukey's biweight estimate of location, the biweight estimate of ability responds sensitively to information from items in the neighborhood of the subject's apparent ability, while effectively discounting responses to items far above or below this level. The response of Subject i to Item j is

assigned the weight W_{ij} in accordance with the distance of the subject from the item, in units of the item's dispersion:

$$W_{ij} = \begin{cases} (1 - U_{ij}^2)^2 & \text{if } |U_{ij}| < 1 \\ 0 & \text{otherwise} \end{cases}$$

with

$$U_{ij} = \frac{b_j - \hat{\theta}_i}{3\, s_j} .$$

In this last expression, $\hat{\theta}_i$ represents the biweight estimate of the ability of Subject i. (The biweight estimate depends on the weights and the weights depend on the estimate; together they must be computed iteratively.)

The fitting function used in the computation of the biweight estimate is a modification of (B.3), the likelihood equation:

$$G' = \sum_{j=1}^{n} W_{ij} \, (X_{ij} - P_{ij})/s_j .$$

As an indication of the precision of estimation, one may use the negative reciprocal of the square root of a facsimile of a second derivative:

$$G'' = - \sum_{j=1}^{n} W_{ij} \, P_{ij} \, (1-P_{ij}) / s_j^2 .$$

This quantity would be the second derivative of a log likelihood if the weights W_{ij} had been specified in advance rather than in response to the data.

BIWEIGHTED BAYES ESTIMATES

Final estimates of subject abilities in the ASVAB power tests use both the standard normal prior, to provide lower mean-squared errors, and biweighting, to trim potentially misleading responses to extreme items. The fitting function incorporates aspects of the Bayes and the biweight estimates:

$$H' = \theta_i + \sum_{j=1}^{n} W_{ij} (X_{ij}-P_{ij})/s_j ,$$

where W_{ij} is the biweight. As a standard error of estimation, we use the negative reciprocal of the square root of a facsimile of a a second derivative:

$$H'' = -1 - \sum_{j=1}^{n} W_{ij} P_{ij} (1-P_{ij})^2 / s_j .$$

THE PROFILE OF AMERICAN YOUTH:

Annotated Bibliography of DoD Related Publications

Compiled by

Brian K. Waters
Human Resources Research Organization

MARCH 1982

Technical Memorandum 82-1
Directorate for Accession Policy
Office of the Secretary of Defense

The Profile of American Youth
Annotated Bibliography of DoD Related Publications

Department of Defense Publications

Department of Defense. Profile of American Youth: 1980 Nationwide Administration of the Armed Services Vocational Aptitude Battery. Washington, D.C.: Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics), March 1982.

In 1980, the Department of Defense (DoD) and the Military Services, in cooperation with the Department of Labor, sponsored a large-scale research project to assess the vocational aptitudes of American youth. A national probability sample of approximately 12,000 young men and women, consisting of participants in the National Longitudinal Survey (NLS) of Youth Labor Force Behavior, were administered the Armed Services Vocational Aptitude Battery (ASVAB).

The young people tested were representative of all youth in the United States, ages 16 through 23 years old. The analyses conducted in the profile study focused upon young people who were 18 through 23 at the time of ASVAB testing. The ASVAB is used by the Military Services to determine enlistment eligibility and assignment qualifications to specific military jobs. Four ASVAB subtests are combined to form the Armed Forces Qualification Test (AFQT), a general measure of trainability and a primary criterion of enlistment eligibility. AFQT scores, reading grade level, and vocational aptitude composite scores were used as indices for comparing the performance of civilian and military groups. The analyses included comparisons of the 1980 youth population with the World War II reference population and with military recruits, as well as comparisons of subgroups within the youth population on the basis of age, sex, race/ethnicity, level of education, socioeconomic status, and geographic region.

The report contains five sections and three appendices, as well as a comprehensive bibliography of relevant literature. Section 1 presents a brief introduction. The study design, sampling procedures, and data analysis are described in Section 2. Section 3 presents a comparison of characteristics of the 1980 youth population with military personnel. In Section 4, the mean scores of the 1980 youth population subgroups are compared. Section 5 summarizes the results documented in the report. The appendices contain detailed statistical data. An executive summary is included at the front of the report.

Eitelberg, M.J. Subpopulation Differences in Performance on Tests of Mental Ability: Historical Review and Annotated Bibliography (Technical Memorandum 81-3). Washington, D.C.: Office of the Secretary of Defense, August 1981.

This paper presents a review of subpopulation differences in performance on tests of mental ability. Six categories of subpopulation differences are examined: sex, age, race, ethnicity, social class, and other (including education, geographical location, and physical characteristics). Selected references (40) are briefly described in an annotated bibliography in order to: (a) acquaint the reader with the so-called "classic" works and a representative sample of studies in the field; (b) characterize the literature dealing with controversial aspects of group differences and intelligence testing; and (c) provide a general "road map" for those who wish to further pursue the subject.

It should be noted that no attempt is made to present, explain, or analyze the possible causes of the observed differences. However, the interested reader can find references to several recent books on causative factors in the annotated bibliography and supplementary notes.

Sellman, W.S., & Hagan, H.T. The Profile of American Youth: Data Audit (Technical Memorandum 81-1). Washington, D.C.: Directorate for Accession Policy, Office of the Secretary of Defense, April 1981.

The Profile of American Youth is a large-scale research project sponsored by the Department of Defense with the cooperation of the Department of Labor. It was designed to assess the vocational aptitudes of contemporary American young people and, at the same time, to establish current national norms for the Armed Services Vocational Aptitude Battery (ASVAB). To achieve these goals, the National Opinion Research Center (NORC) of the University of Chicago administered the ASVAB during the summer and fall of 1980 to a national probability sample of nearly 12,000 men and women ages 16 to 23. Because of the importance of this study, both to the Department of Defense and to the social science community, it was imperative that an independent audit of the test scores, the demographic variables, and the procedures used to assemble the data base be conducted.

A Department of Defense (DoD) team of computer programmers, testing and survey experts performed a data audit to ensure the accuracy of the information contained on the data tape to be provided by NORC to DoD. This report documents the procedures used in the audit and presents its results. The report addresses several issues: (a) it summarizes the research questions that motivated the aptitude profile study, (b) it describes the sample to which the ASVAB was administered, (c) it explains

the various sources of data assembled in the analytic data base, and (d) it reports on the activities performed by the DoD audit team to verify and audit the data base.

Waters, B.K. The Test Score Decline: A Review and Annotated Bibliography (Technical Memorandum 81-2). Washington, D.C.: Directorate for Accession Policy, Office of the Secretary of Defense, August 1981.

This paper brings together the large body of literature on indicators of changes in U.S. scholastic aptitude and achievement levels within the 1952-1980 period. The target population is youth from pre-school age to post-college graduate age. The study begins with a discussion of the trends which have emerged over time, supported by credible data. The discussion reviews aptitude test score patterns from 1952 through 1980 on eight scholastic achievement tests. An annotated bibliography of 49 sources was reviewed by the author. A bibliography of literature related to the decline of tests scores is included. This reference list includes over 250 sources.

The aptitude testing data covering the entire period from the early 1950s to 1980 show remarkable consistency. With the exception of slight increases on the LSAT, the other measures of scholastic aptitude of youth reviewed have consistently decreased at a rate of about 1 to 3 percent of a standard deviation per year. Although there is some evidence that the rate of decline has lessened somewhat in the past three years, the trend continues through 1980.

- In general, the author also found consistent evidence of achievement test score declines, for the 1960s through 1970s in all areas tested above grade 4. The author concluded that these trends are real, national in scope, and continuing, though at a decreasing rate of decline since about 1977.

National Opinion Research Center Publications

Bock, R.D., & Mislevy, R.J. Data Quality Analysis of the Armed Services Vocational Aptitude Battery. Chicago: National Opinion Research Center, August 1981.

The Profile of American Youth is a large-scale research project sponsored by the Department of Defense with cooperation of the Department of Labor. It assessed the vocational aptitudes of contemporary American young people and, at the same time, established current national norms for the Armed

Services Vocational Aptitude Battery (ASVAB). To achieve these goals, the National Opinion Research Center (NORC) administered ASVAB Form 8A during the summer and fall of 1980 to a national probability sample of nearly 12,000 men and women, ages 16 to 23.

Because of the importance of this study, not only to its sponsors but to the social science research community and the general public, care was taken to examine and document the quality of the data collected. This review included the ASVAB's suitability for assessing the aptitudes of the 1980 youth population. Whether the data collected are appropriate for this purpose depends on (a) the adequacy of the sampling plan and its implementation; (b) the quality of the fieldwork and test administration procedures; and (c) the psychometric quality of the test data collected. This report examines the test data.

Using item response theory (IRT) methods, the authors investigated (a) item response profiles of individual subjects for evidence of unusual patterns of right and wrong answers which could indicate carelessness, malingering, or guessing; (b) the possibility of test item bias that could favor one or another subcultural group; (c) the reliability of the ASVAB subtests and the amount of information they provide about subjects across the entire range of ability; (d) the consistency of test administration at the several hundred testing centers established across the country.

The authors concluded that the ASVAB is useful for measuring vocational aptitudes of civilian youth. They found the ASVAB free from major defects such as high levels of guessing or carelessness, inappropriate levels of difficulty, cultural test-question bias, and inconsistencies in test administration procedures.

Bock, R.D., & Moore, E.G.J. The Profile of American Youth: Demographic Influences on ASVAB Test Performance. Chicago: National Opinion Research Center, December 1981.

In the summer and fall of 1980, the National Opinion Research Center of the University of Chicago administered the Department of Defense enlistment test, the Armed Services Vocational Aptitude Battery (ASVAB), to a national probability sample of approximately 12,000 young men and women between the ages of 16 and 23. The ASVAB is currently used, in conjunction with other entry standards, to select personnel for the Army, Navy, Marine Corps, and Air Force, and to assign them to appropriate military jobs. The battery consists of ten tests, which measure knowledge and skill in these areas: General Science, Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Numerical Operations, Coding Speed, Auto and Shop Information, Mathematics Knowledge, Mechanical Comprehension, and Electronics Information.

The ASVAB was administered to this sample to obtain current national percentile norms for the tests and to assess the vocational potential of the contemporary youth population. Because the subjects used in this study had previously been interviewed in connection with the Department of Labor's National Longitudinal Survey of Youth Labor Force Behavior, considerable demographic background information was available on each subject in addition to the aptitude test score information. This combination of test score and background data provide a virtually unique opportunity to describe in detail the vocational aptitude test performance of a truly representative sample of American young people.

Of the various background factors covered in the interview, the following show notable effects on average test score profiles: (a) sex, (b) highest grade completed, (c) sociocultural group, (d) economic status, (e) region of residence at age 14, and (f) mother's highest grade completed. Effects of these factors on the ten ASVAB subtests were studied by multivariate analysis of variance. The background variables used in the analysis were defined as follows: sex--male, female; highest grade completed--0-8, 9-11, 12, 13+; sociocultural group--white, black, Hispanic; economic status--non-poor, poor (below poverty level); region of residence at age 14--Northeast, Southeast, Midwest, West; mother's highest grade completed--0-8, 9-11, 12, some college, college graduate.

The analysis showed statistically significant effects or interactions of these background factors on one or more of the ASVAB subtests. Three-factor interactions were observed for: sex by sociocultural group by economic status; highest grade completed by sociocultural group by economic status; and sociocultural group by region of residence at age 14 by economic status. A significant two-factor interaction not represented in the triple interactions was sex by highest grade completed. A main effect not represented in any interaction was mother's highest grade completed.

Frankel, M.R. & McWilliams, H.A. The profile of American youth: Technical sampling report. Chicago: National Opinion Research Center, March 1981.

The Profile of American Youth is a large-scale research project sponsored by the Department of Defense (DoD) with the cooperation of the Department of Labor. Its purposes were to assess the vocational aptitudes of contemporary American young people and, at the same time, establish current national norms for the Armed Services Vocational Aptitude Battery (ASVAB). To achieve these goals, the National Opinion Research Center administered the ASVAB during the summer and fall of 1980 to a national probability

sample of nearly 12,000 young men and women. Because of the importance of the sampling components of the Profile study to the overall reliability of the ability estimates produced, a peer review committee of experts in the field of survey sampling was recruited to provide an independent review of the sample design and estimation procedures used in the study.

The overall study design for the Profile of American Youth required that the ASVAB be administered under standard conditions to a nationally representative sample of young people. Rather than selecting an entirely new national probability sample, a lengthy and expensive task, NORC and the sponsors of the study jointly decided to utilize the already existing sample of youth initially selected for the five-year National Longitudinal Survey (NLS) of Labor Force Behavior. This sample, selected in 1978 and interviewed in 1979 and 1980, was a nationally representative sample of American youth ages 16 to 23 (born 1957 to 1964) and, therefore, met the requirements for the Profile of American Youth. Moreover, both the Department of Defense, as a cosponsor of the NLS (with the Department of Labor), and NORC, as the NLS subcontractor for data collection, were already familiar with the sample.

It is important to emphasize that the profile sample is built on the NLS sample. The profile study began with the list of respondents who had completed the baseyear (1979) interview for the NLS. This was the "target group" to which the ASVAB was administered.

This report provides a detailed, technical discussion of the design of the sample, selection and implementation of the sample, sample weighting, the reliability of the results, and an evaluation of the potential impact of nonresponse on study results. A non-technical version of this report (McWilliams & Frankel, 1981) is also available for the less technically-oriented reader (see below).

McWilliams, H.A. The Profile of American Youth: Field Report. Chicago: National Opinion Research Center, December 1980.

The Profile of American Youth is a large-scale research project sponsored by the Department of Defense with the cooperation of the Department ofLabor. It was designed to assess the vocational aptitudes of contemporary American young people and, at the same time, establish current national norms for the Armed Services Vocational Aptitude Battery (ASVAB). To achieve these goals, the National Opinion Research Center administered theASVAB during the summer and fall of 1980 to a national probability sample of nearly 12,000 men and women ages 16 to 23. This field report addresses several issues: it discusses the research questions that motivated the study; it illustrates the planning and strategic

considerations that went into the design of the project; it describes the sample to which the test was administered; it explains the organization and management of the data collection effort, including (a) the field reporting structure, (b) the training provided to the field staff, (c) the recruitment of respondents, (d) the test administration procedures, and (e) the processing of the data collected. It concludes by presenting the results of the data collection effort in terms of the participation rate and by discussing the respondents' reactions to their participation in the study. Throughout the report an attempt is made to explain not just what was done, but why it was done as well.

McWilliams, H.A., & Frankel, M.R. The Profile of American Youth: Non-Technical Sampling Report. Chicago: National Opinion Research Center. January 1982.

The Profile of American Youth is a large-scale research project sponsored by the Department of Defense (DoD) with the cooperation of the Department of Labor. Its purposes were to assess the vocational aptitudes of contemporary American young people and, at the same time, establish current national norms for the Armed Services Vocational Aptitude Battery (ASVAB). To achieve these goals, the National Opinion Research Center admininstered the ASVAB during the summer and fall of 1980 to a national probability sample of nearly 12,000 young men and women.

This report is a non-technically-oriented version of the technical sampling report (Frankel & McWilliams, 1981) cited above. The report describes the design, selection, implementation, and limited statistical description of the study sample. This report is designed for the reader interested in knowing what was done and how, rather than the detailed statistical presentations contained in its companion report. It is written in language understood by the non-statistician.

Sheatsley, P.B. The Profile of Americn Youth: Pretest Report. Chicago: National Opinion Research Center, September 1980.

The Profile of American Youth is a large-scale research project sponsored by the Department of Defense (DoD) with the cooperation of the Department of Labor. It was designed to assess the vocational aptitudes of contemporary American young people and, at the same time, establish current national norms for the Armed Serices Vocational Aptitude Battery (ASVAB).

To achieve these goals, the National Opinion Research Center (NORC) administered the ASVAB during the summer and fall of 1980 to a national probability sample of nearly 12,000 men and women aged 16 to 23.

The study had its genesis in the National Longitudinal Study (NLS) of Youth Labor Force Behavior, which is funded by the Department of Labor under authority of the Comprehensive Employment and Training Act. Supplementary funding for NLS has been provided by the Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics) and the Military Services.

The purpose of the NLS is to study a large and representative cross-section of American youth through annual personal interviews over a five-year period. Young people are at a crucial point in their lives, a time when they make decisions about education, career, and family formation. While a central concern of the study is problems relating to employment and unemployment, the interviews gather a great deal of supplemental information about the characteristics, experiences, plans, and attitudes of the youth.

Although NORC already had access to the NLS panel of young people, there still remained the formidable problems of persuading them to take ASVAB, selecting and preparing suitable test locations, and training NORC interviewers to administer the test under a wide variety of circumstances. For these reasons, NORC formally proposed to DoD that a pretest of the profile study be conducted in the spring of 1980. This pretest provided a tryout of the proposed materials and procedures in a small number of places to test their reliability, feasibility, and cost implications before proceeding with the full study.

The pretest was extremely useful in providing an empirical test of NORC's methodology, and the lessons learned were adapted to the full-scale testing during July through October 1980. The present report provides a detailed account of pretest activities.

NLSY79
PROFILES OF AMERICAN YOUTH

ADDENDUM to ATTACHMENT 106

April 1992

NLS User Services
usersvc@postoffice.chrr.ohio-state.edu

ATTACHMENT 106 ADDENDUM

There were 13 new variables added to the NLSY record type PROFILES for the 1979-1990 release. These include: (1) ten subtest standard scores, R(6180.10) to R(6180.19), one for each of the 10 subtests raw scores; (2) a standard score for a composite variable called Verbal, R(6181.); (3) and two Armed Forces Qualification Test (AFQT) percentile scores, R(6182.) and R(6183.). PROFILES variables available prior to the 1979-1990 release include subtest raw scores (i.e. the number of correct answers), scale scores, and standard errors.

All 12,686 NLSY respondents were eligible to participate in the PROFILES study; 11,914 respondents completed the tests. The standard scores and percentile scores are computed for all valid raw subtest scores, except for thirty-six respondents who were deemed to have taken the test under altered test conditions (code 67 on R(6148.)). Those respondents with altered test procedures have been assigned a -3 value on the standard scores and percentile scores. It should be noted that the standard scores and percentile scores are based on those respondents born prior to 1963 on R(5.) and who do not have altered test procedures.

Note: Tables A through E referenced within the text of this addendum appear at the end of this attachment.

PROFILES Subtest Standard Scores

10 Subtest Standard Scores - R(6180.10) to R(6180.19)

Each of the ten PROFILES subtest raw scores were converted (recoded) to subtest standard scores by using the figures in Tables C and D.

Verbal (VE) Composite Standard Score - R(6181.)

The Verbal (VE) composite was computed by first summing the raw scores for Word Knowledge and Paragraph Comprehension (WK + PC = VE) and then converting to a standard score by using the figures in Table D.

The administration of the PROFILES tests involved the use of answer sheets that differed from the answer sheets used by the Department of Defense (DoD) in the administration of the Armed Services Vocational Aptitude Battery (ASVAB). The inconsistency has resulted in the need to make small adjustments to the PROFILES speeded subtests of Numerical Operations (NO) and Coding Speed (CS). The corrections estimate the raw score a respondent would have obtained if the military answer sheet had been used.

The figures in Table A provide the crosswalk used to make the needed adjustments to the PROFILES Numerical Operations and Coding Speed raw scores. The PROFILES raw scores (unadjusted) for Numerical Operations and Coding Speed, R(6154.) and R(6155.) respectively, were adjusted using the figures in Table A. These adjusted

raw scores were then used in the conversion to standard scores for Numerical Operations and Coding Speed, R(6180.14) and R(6180.15) respectively. The other tests were not affected by the answer sheet differences.

The standard scores are linear transformations of raw scores and, when weighted by R(6147.) for those respondents born prior to 1963 with unaltered test procedures, exhibit a mean of 50 and a standard deviation of 10, approximately.

PROFILES Armed Forces Qualification Test (AFQT)

AFQT Definition Through 1988 - R(6182.)

Two Armed Forces Qualification Test (AFQT) percentile scores were added to the Main NLSY data file for the 1979-1990 release. The formula to compute R(6182.) was used by the DoD through 1988 (some NLSY documents will refer to this variable as AFQT80). It was computed as follows:

1. Adjust the Numerical Operations (NO) raw score for the answer sheet differences using the figures in Table A.

2. Compute an AFQT80 "raw" score as the sum of Word Knowledge (WK), Paragraph Comprehension (PC), Arithmetic Reasoning (AR), and one-half of the Numerical Operations (NO) raw scores.

   WK + PC + AR + (.5(NO)) = AFQT80 raw score

3. Convert the AFQT80 raw score to a percentile score using the figures in Table B.

AFQT Definition Since January, 1989 - R(6182.)

The formula to compute R(6183.) has been used operationally since January, 1989 (some NLSY documents will refer to this variable as AFQT89). It was computed as follows:

1. Convert the raw scores for Arithmetic Reasoning (AR), Math Knowledge (MK), and Verbal (VE) composite to standard scores using the figures in Tables C and D.

2. Compute a sum of standard scores by adding the standard scores for AR and MK and two times the VE.

3. Convert the AFQT89 standard score to a percentile score using the figures in Table E.

The AFQT80 and AFQT89 percentile scores, when weighted by R(6147.) for those respondents born prior to 1963 with unaltered test procedures, exhibit a mean of 50 and a standard deviation of 29, approximately.

TABLE A

ADJUSTMENT TO NUMERICAL OPERATIONS AND CODING SPEED
RAW SCORES IN THE 1980 YOUTH POPULATION

Original score in 1980 Youth Population	Adjusted score		Score in 1980 Youth Population	Adjusted score	
	Numerical Operations	Coding Speed		Numerical Operations	Coding Speed
0	0	0	43	46	44
1	0	1	44	47	45
2	1	2	45	48	46
3	2	3	46	49	47
4	4	4	47	49	48
5	5	5	48	50	49
6	6	6	49	50	50
7	8	7	50	50	51
8	9	8	51		53
9	10	9	52		54
10	11	10	53		55
11	12	11	54		56
12	14	12	55		57
13	15	13	56		58
14	16	14	57		59
15	17	15	58		60
16	18	16	59		61
17	19	17	60		62
18	21	18	61		63
19	22	20	62		64
20	23	21	63		65
21	24	22	64		66
22	25	23	65		67
23	26	24	66		68
24	27	25	67		69
25	28	26	68		70
26	29	27	69		71
27	30	28	70		72
28	31	29	71		73
29	33	30	72		74
30	34	31	73		75
31	35	32	74		76
32	36	33	75		77
33	37	34	76		78
34	38	35	77		79
35	39	36	78		80
36	39	37	79		81
37	40	38	80		82
38	41	39	81		83
39	42	40	82		84
40	43	41	83		84
41	44	42	84		84
42	45	43			

Source: Air Force Human Resources Laboratory, TR-85-14, *The 1980 Youth Population: Correcting the Speeded Tests*, by T.G. Wagner and M.J. Ree, July 1985.

TABLE B

ARMED FORCES QUALIFICATION TEST (AFQT)
1980 PERCENTILE EQUIVALENTS (Prior to 1 Jan 89)

RAW AFQT SCORE	PERCENTILE	RAW AFQT SCORE	PERCENTILE	RAW AQFT SCORE	PERCENTILE	RAW AFQT SCORE	PERCENTILE	RAW AFQT SCORE	PERCENTILE
0.0	1	21.5	1	43.0	11	64.5	30	86.0	67
0.5	1	22.0	1	43.5	11	65.0	30	86.5	68
1.0	1	22.5	1	44.0	11	65.5	31	87.0	69
1.5	1	23.0	1	44.5	12	66.0	32	87.5	70
2.0	1	23.5	1	45.0	12	66.5	32	88.0	71
2.5	1	24.0	2	45.5	12	67.0	33	88.5	72
3.0	1	24.5	2	46.0	13	67.5	34	89.0	73
3.5	1	25.0	2	46.5	13	68.0	35	89.5	74
4.0	1	25.5	2	47.0	13	68.5	35	90.0	75
4.5	1	26.0	2	47.5	14	69.0	36	90.5	76
5.0	1	26.5	2	48.0	14	69.5	37	91.0	77
5.5	1	27.0	2	48.5	14	70.0	38	91.5	78
6.0	1	27.5	3	49.0	15	70.5	38	92.0	79
6.5	1	28.0	3	49.4	15	71.0	39	92.5	80
7.0	1	28.5	3	50.0	16	71.5	40	93.0	81
7.5	1	29.0	3	50.5	16	72.0	41	93.5	82
8.0	1	29.5	3	51.0	16	72.5	42	94.0	83
8.5	1	30.0	4	51.5	17	73.0	42	94.5	84
9.0	1	30.5	4	52.0	17	73.5	43	95.0	85
9.5	1	31.0	4	52.5	17	74.0	44	95.5	86
10.0	1	31.5	4	53.0	18	74.5	45	96.0	87
10.5	1	32.0	4	53.5	18	75.0	46	96.5	88
11.0	1	32.5	5	54.0	19	75.5	46	97.0	89
11.5	1	33.0	5	54.5	19	76.0	47	97.5	90
12.0	1	33.5	5	55.0	20	76.5	48	98.0	91
12.5	1	34.0	5	55.5	20	77.0	49	98.5	92
13.0	1	34.5	6	56.0	21	77.5	49	99.0	93
13.5	1	35.0	6	56.5	21	78.0	50	99.5	94
14.0	1	35.5	6	57.0	22	78.5	51	100.0	94
14.5	1	36.0	6	57.5	22	79.0	52	100.5	95
15.0	1	36.5	6	58.0	23	79.5	53	101.0	96
15.5	1	37.0	7	58.5	23	80.0	54	101.5	97
16.0	1	37.5	7	59.0	24	80.5	55	102.0	98
16.5	1	38.0	7	59.5	24	81.0	56	102.5	98
17.0	1	38.5	8	60.0	25	81.5	57	103.0	99
17.5	1	39.0	8	60.5	25	82.0	58	103.5	99
18.0	1	39.5	8	61.0	26	82.5	59	104.0	99
18.5	1	40.0	8	61.5	26	83.0	60	104.5	99
19.0	1	40.5	9	62.0	27	83.5	62	105.0	99
19.5	1	41.0	9	62.5	27	84.0	63		
20.0	1	41.5	10	63.0	28	84.5	64		
20.5	1	42.0	10	63.5	28	85.0	65		
21.0	1	42.5	10	64.0	29	85.5	66		

NOTE: NOT APPLICABLE FOR ASVAB FORMS 15/16/17 (EXCEPT FORM 15c)

SOURCE: Department of Defense (Oct. 90).
Conversion Tables: Armed Services Vocational Aptitude Battery (ASVAB) Forms 8-19. (D.D 1304,12W1).

TABLE C

CONVERSION OF RAW TEST SCORES TO 1980 STANDARD SCORE EQUIVALENTS

RAW	GS	AR	WK	PC	NO	CS	RAW	RAW	GS	AR	WK	PC	NO	CS	RAW
0	20	26	20	20	20	22	0	45					57	48	45
1	20	27	20	20	20	22	1	46					58	49	46
2	22	28	20	23	20	23	2	47					59	50	47
3	24	30	20	26	20	23	3	48					60	50	48
4	26	31	21	29	20	24	4	49					61	51	49
5	28	32	22	32	20	25	5	50					62	51	50
6	30	34	24	35	21	25	6	51						52	51
7	32	35	25	38	22	26	7	52						53	52
8	34	36	26	41	23	26	8	53						53	53
9	36	38	28	44	24	27	9	54						54	54
10	38	39	29	47	25	28	10	55						54	55
11	40	40	30	50	26	28	11	56						55	56
12	42	42	31	53	27	29	12	57						56	57
13	44	43	33	56	28	29	13	58						56	58
14	46	45	34	59	28	30	14	59						57	59
15	48	46	35	62	29	31	15	60						57	60
16	50	47	37		30	31	16	61						58	61
17	52	49	38		31	32	17	62						59	62
18	54	50	39		32	32	18	63						59	63
19	56	51	41		33	33	19	64						60	64
20	58	53	42		34	34	20	65						60	65
21	60	54	43		35	34	21	66						61	66
22	62	55	44		36	35	22	67						62	67
23	64	57	46		37	35	23	68						62	68
24	66	58	47		38	36	24	69						63	69
25	68	59	48		39	37	25	70						63	70
26		61	50		40	37	26	71						64	71
27		62	51		41	38	27	72						65	72
28		64	52		41	38	28	73						65	73
29		65	54		42	39	29	74						66	74
30		66	55		43	39	30	75						66	75
31			56		44	40	31	76						67	76
32			57		45	41	32	77						68	77
33			59		46	41	33	78						68	78
34			60		47	42	34	79						69	79
35			61		48	42	35	80						69	80
36					49	43	36	81						70	81
37					50	44	37	82						71	82
38					51	44	38	83						71	83
39					52	45	39	84						72	84
40					53	45	40								
41					53	46	41								
42					54	47	42								
43					55	47	43								
44					56	48	44								

SOURCE: Department of Defense (Oct. 90).
Conversion Tables: Armed Services Vocational Aptitude Battery (ASVAB) Forms 8-19. (D.D 1304,12W1).

TABLE D

CONVERSION OF RAW TEST SCORES TO 1980 STANDARD SCORE EQUIVALENTS

RAW	AS	MK	MC	EI	VE	RAW	RAW	AS	MK	MC	EI	VE	RAW
0	24	29	24	23	20	0	25	69	68	70		38	25
1	26	30	25	25	20	1	26					39	26
2	28	32	27	27	20	2	27					40	27
3	30	33	29	30	20	3	28					41	28
4	31	35	31	32	20	4	29					42	29
5	33	37	33	34	20	5	30					43	30
6	35	38	35	37	20	6	31					44	31
7	37	40	37	39	21	7	32					45	32
8	39	41	38	42	22	8	33					46	33
9	40	43	40	44	23	9	34					47	34
10	42	44	42	46	24	10	35					48	35
11	44	46	44	49	25	11	36					49	36
12	46	48	46	51	26	12	37					50	37
13	48	49	48	53	27	13	38					51	38
14	49	51	50	56	28	14	39					52	39
15	51	52	52	58	29	15	40					53	40
16	53	54	53	60	30	16	41					54	41
17	55	55	55	63	31	17	42					54	42
18	57	57	57	65	32	18	43					55	43
19	58	58	59	68	33	19	44					56	44
20	60	60	61	70	34	20	45					57	45
21	62	62	63		35	21	46					58	46
22	64	63	65		36	22	47					59	47
23	66	65	67		37	23	48					60	48
24	67	66	68		37	24	49					61	49
							50					62	50

SOURCE: Department of Defense (Oct. 90).
Conversion Tables: Armed Services Vocational Aptitude Battery (ASVAB) Forms 8-19. (D.D 1304,12W1).

TABLE E

ARMED FORCES QUALIFICATION TEST (AFQT = 2VE + AR + MK)
STANDARD SCORE TO 1980 PERCENTILE CONVERSIONS
(After 1 Jan 89)

STD Score	Percentile	STD Score	Percentile	STD Score	Percentile	STD Score	Percentile
80 to 95	1	140	8	185	33	230	76
96	1	141	8	186	34	231	77
97	1	142	8	187	35	232	78
98	1	143	9	188	35	233	79
99	1	144	9	189	36	234	80
100	1	145	10	190	37	235	80
101	1	146	10	191	38	236	81
102	1	147	11	192	39	237	82
103	1	148	11	193	40	238	84
104	1	149	12	194	41	239	84
105	1	150	12	195	42	240	85
106	1	151	13	196	42	241	86
107	1	152	13	197	43	242	87
108	1	153	13	198	44	243	88
109	1	154	14	199	45	244	89
110	1	155	15	200	46	245	90
111	1	156	15	201	47	246	91
112	1	157	16	202	48	247	92
113	1	158	16	203	49	248	93
114	1	159	17	204	50	249	94
115	1	160	17	205	51	250	95
116	1	161	18	206	52	251	96
117	1	162	18	207	53	252	97
118	1	163	19	208	53	253	98
119	1	164	19	209	54	254	99
120	1	165	20	210	55	255	99
121	2	166	21	211	56	256	99
122	2	167	21	212	57	257	99
123	2	168	22	213	58	258 to 320	99
124	2	169	22	214	59		
125	3	170	23	215	61		
126	3	171	23	216	62		
127	3	172	24	217	63		
128	4	173	25	218	64		
129	4	174	25	219	65		
130	4	175	26	220	66		
131	4	176	27	221	67		
132	5	177	27	222	68		
133	5	178	28	223	69		
134	5	179	29	224	70		
135	6	180	29	225	71		
136	6	181	30	226	72		
137	6	182	31	227	73		
138	7	183	32	228	74		
139	7	184	32	229	75		

SOURCE: Department of Defense (Oct. 90).
Conversion Tables: Armed Services Vocational Aptitude Battery (ASVAB) Forms 8-19. (D.D 1304,12W1).
[doty.nlsy-doc-1992]profiles.addendum

References

BOCK, R. DARRELL and MOORE, ELSIE G.J. *Advantage and Disadvantage: A Profile of American Youth.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1986.

CLAUDY, JOHN G. and STEEL, LAURI. "Armed Services Vocational Battery: Validation for Civilian Occupations Using National Longitudinal Survey of Youth." AFHRL Technical Report 90-29, American Institute for Research, 1990.

FRANKEL, MARTIN R. and McWILLIAMS, HAROLD A. "The Profile of American Youth: Technical Sampling Report." Chicago: NORC,1981.

McWILLIAMS, HAROLD A. and FRANKEL, MARTIN R. "The Profile of American Youth: Non-Technical Sampling Report." Chicago: NORC, 1982.

McWILLIAMS, HAROLD A. "The Profile of American Youth: Field Report." Chicago, NORC, 1980.

WEGNER, T.G. and REE, M.J. "The 1980 Youth Population: Correcting the Speeded Tests." TR-85-14, Air Force Human Resources Laboratory, July 1985