HumRRO

FR-01-12

CONTRACT FOR MANPOWER, PERSONNEL, AND TRAINING ANALYSES (COMPTA II) FOR THE DEFENSE MANPOWER DATA CENTER (DMDC)

Testing Via the Internet: A Literature Review and Analysis of Issues for Department of Defense Internet Testing of the Armed Services Vocational Aptitude Battery (ASVAB) in High Schools

James R. McBride
Arthur F. Paddock
Lauress L. Wise
William J. Strickland
Brian K. Waters

Submitted To:

Defense Manpower Data Center
400 Gigling Road
Seaside, CA 93955-6771

Prepared under:

Contract Number M67004-00-D0002
Delivery Order 0007

May 11, 2001

Human Resources Research Organization
66 Canal Center Plaza, Suite 400 • Alexandria, VA 22314-1591
703-549-3611 • www.HumRRO.org

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy)	2. REPORT TYPE	3. DATES COVERED (from. . . to)
May 11, 2001	Final Report	July 14, 2000 – December 31, 2000

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
Testing Via the Internet: A Literature Review and Analysis of Issues for DoD Internet Testing of the ASVAB in High Schools	M67004-00-D-0002 (Delivery Order 0007)
	5b. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) James R. McBride, Arthur F. Paddock, Lauress L. Wise, William J. Strickland, and Brian K. Waters	5c. PROJECT NUMBER
	5d. TASK NUMBER
	5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Human Resources Research Organization 66 Canal Center Plaza, Ste 400 Alexandria, VA 22314	8. PERFORMING ORGANIZATION REPORT NUMBER FR-01-12
SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Defense Manpower Data Center 400 Gigling Road Seaside, CA 93955	10. MONITOR ACRONYM DMDC
	11. MONITOR REPORT NUMBER

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Dr. John Welsh, Contracting Officer's Representative

14. ABSTRACT *(Maximum 200 words)*:

This report identified key technical issues involved with the presentation of CAT-ASVAB over the Internet to high schools, and assessed the feasibility of developing an alternative test administration system that could simultaneously serve the goals of providing career exploration information to high schools and providing recruiters with qualified lead lists. Testing/measurement and industrial/organizational psychological literature were surveyed. Organizations that use Internet—delivered employment testing or employment screening were contacted, identifying major issues associated with establishment and operation of such systems. An ordered list of options for prototype development were rank-ordered from high stakes/high risk scenarios to low stakes, low risk scenarios based on knowledge gained.

15. SUBJECT TERMS

Personnel, Testing, ASVAB, Internet

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited	TBD	Dr. John Welsh (831) 583-4171

Standard Form 298

FR-01-12

CONTRACT FOR MANPOWER, PERSONNEL, AND TRAINING ANALYSES (COMPTA II) FOR THE DEFENSE MANPOWER DATA CENTER (DMDC)

Testing Via the Internet: A Literature Review and Analysis of Issues for Department of Defense Internet Testing of the Armed Services Vocational Aptitude Battery (ASVAB) in High Schools

James R. McBride
Arthur F. Paddock
Lauress L. Wise
William J. Strickland
Brian K. Waters

Submitted To:

Defense Manpower Data Center
400 Gigling Road
Seaside, CA 93955-6771

Prepared under:

Contract Number M67004-00-D0002
Delivery Order 0007

May 11, 2001

Human Resources Research Organization
66 Canal Center Plaza, Suite 400 • Alexandria, VA 22314-1591
703-549-3611 • www.HumRRO.org

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

Study Overview[1]

The purpose of this study was to identify the key technical issues involved with the presentation of the Armed Services Vocational Aptitude Battery (ASVAB) via the Internet in the Department of Defense Student Testing Program (STP), and to determine the feasibility of developing an alternative test administration system that could simultaneously serve the goals of providing career exploration information to high schools and to provide recruiters with qualified lead lists.

The study consisted of four major tasks:

1. Survey the testing and measurement literature, and the industrial and organizational psychological literature and survey agencies and organizations that may use Internet-delivered employment testing, or employment screening.

2. Identify major issues surrounding the use of the Internet for employment testing, career exploration, or employment screening.

3. Provide an ordered list of options for prototype development rank ordered from high stakes / high risk scenarios to low stakes/ low risk scenarios based on knowledge gained.

4. Identify a milestone schedule for each option and provide a detailed listing of tasks necessary to accomplish the implementation of options selected schedule by the monitoring agency.

This report is organized into four parts, corresponding to those four tasks.

CAT-ASVAB Background Information

ASVAB is currently delivered via two media. The first is paper-and-pencil; the second is by computer. Where administered by computer, it is also adaptive (See Chapter 1; Sands, Waters, & McBride, 1997). Adaptive versions of ASVAB are referred to collectively as Computer Adaptive Testing – ASVAB (CAT-ASVAB). CAT-ASVAB is currently used only in the Enlistment Testing Program (ETP). In that program, CAT-ASVAB is used in all of the Military Entrance Processing Stations (MEPS). Except for an experimental trial now in progress, all mobile examining team (MET) sites still use the paper-and-pencil medium. In the Student Testing Program (STP), ASVAB is administered to high school and community college students. All STP testing currently uses the paper-and-pencil medium.

If the Internet were to be used to administer ASVAB in the STP, presumably that would entail the use of CAT-ASVAB. A useful introduction to this study is a brief consideration of some unique features of CAT-ASVAB.

First of all, CAT-ASVAB is adaptive. Test items are chosen one-by-one, contingent on the examinee's own performance, with the aim of matching the difficulty of the test to the ability level of the examinee. This results in substantial measurement efficiency, which is capitalized on to make CAT-ASVAB a

[1] The authors wish to thank Dr. John Welsh, the project monitor, and the Monterey Defense Manpower Data Center staff for their invaluable input to the project and this report. We also are indebted to the professionals who were interviewed provided their expertise to the information gathering phase of the work.

much shorter test battery than its printed counterparts—about 100 minutes in contrast to over three hours for the paper-and-pencil version.

Second, the CAT-ASVAB software system is an MS-DOS application, as opposed to a Microsoft Windows application. It does not make use of the Windows application programming interface, and does not have the now familiar look and feel of software that uses the Windows graphical user interface.

Third, CAT-ASVAB uses a customized keyboard for input of examinee responses. Mouse input is not an option, and the keyboard itself contains only a few keys. Thus, the input medium most familiar to computer-literate people is not used in CAT-ASVAB.

Fourth, all CAT-ASVAB test administration in the ETP takes place on government-owned equipment in government-operated facilities.

Fifth, the controlling software, as well as the ASVAB item banks, is permanently stored in storage media that are part of the local equipment.

Internet / Current CAT-ASVAB Differences

An Internet-administered version of ASVAB for use in the STP might differ from the ETP version of CAT-ASVAB in the other aspects just discussed. These differences can be summarized as differences in equipment ownership, software, user interface, and locus of data storage. Each is discussed briefly below.

Equipment ownership

A distinguishing feature of the STP is that ASVAB is administered in cooperating schools' own facilities, to as many students as the school wishes. If an Internet-based STP delivery system were implemented, it is not likely to be practical for government-owned equipment to be used, unless the government installed such equipment permanently in the affected schools. On the contrary, one of the most attractive features of Internet administration would be that it would provide schools with access to CAT-ASVAB using their own equipment, such as computers in school computer labs. An alternative to the use of school equipment would be to have the test administered by vendors, either in vendor-operated facilities, or in the schools using portable equipment.

Graphical user interface

Although it might be possible for an Internet version of CAT-ASVAB to be designed such that all screen displays were identical (or nearly so) to those used in the current MS-DOS version, it seems more likely that a more modern graphical user interface would be used, one more similar to other Internet applications. Most new computers and operating systems will not use DOS. [2]

Response input medium

With the exception of highly specialized applications (such as Internet kiosks seen in some airports, shopping centers and other public places), virtually all Internet applications use mouse input, which by now has become almost

[2] A Microsoft Windows implementation of CAT-ASVAB is in preparation. CAT-ASVAB will probably be a Windows-based application by the time an Internet version of ASVAB could be implemented.

ubiquitous and highly familiar. An Internet application that did not support mouse input would be unusual, and difficult for many prospective examinees to use.

Testing control software

Although it might be possible for an Internet-delivered version of CAT-ASVAB to use a modified version of the current MS-DOS control software, this seems unlikely for at least two reasons. First, it would limit Internet delivery to computers that are MS-DOS compatible, eliminating most Apple equipment, as well as most equipment using the Unix operating system. Second, it would require MS-DOS to be resident on all computers. This might eliminate even so-called "WINTEL" PCs, as future versions of the Windows operating systems may not support MS-DOS at all.

This suggests one of two approaches to the software used to control an Internet version of ASVAB. The first is to adapt the current system for use with existing Internet testing control software. The second is to develop new CAT-ASVAB software that is Internet-compatible.

Data storage locations

"Data" here refers to three distinct kinds of information:

- The executable program code that controls CAT-ASVAB;
- The ASVAB test item bank databases; and
- The examinee data files (test results) that are produced each time CAT-ASVAB is administered.

In the current CAT-ASVAB system, all these data are stored on the local CAT-ASVAB system. Additionally, extracts of test results files are uploaded to United States Military Entrance Processing Command (USMEPCOM) by means of a communications link between the CAT-ASVAB equipment and another system located in each MEPS.

In an Internet-based system, any or all of these data might be stored at locations remote from the test administration sites. Remote storage typically would be on a network host server. Although continuing to store some or all data locally is an option, considerations such as data security or storage capacity constraints might mitigate towards remote storage of at least some data.

Study Objective A. Survey the testing and measurement and industrial and organizational psychological literature, and survey agencies and organizations that may use Internet—delivered employment testing, or employment screening

For this task, HumRRO conducted surveys of the journal literature in two fields: Tests and measurement, and industrial / organizational psychology. We also conducted Internet searches seeking to identify users of Internet-delivered employment tests and employee screening devices, as well as organizations and companies providing related Internet software and services.

Almost nothing was found in searches of professional journal literature, and searches for companies using Internet-delivered employment tests were disappointing. In contrast, between professional contacts and searches of the world wide web, we identified a number of

products and services of interest, as well as a number of companies offering services in the area of Internet assessment.

This section of the report will focus on summary descriptions of the products and services we found, and on an integration of the myriad of information obtained. There appears to be a growing number of companies offering either test delivery over the Internet, or Internet services related to test delivery. They can be differentiated using the following dimensions: 1) The assessment market they serve, and 2) The nature of the services offered. For the most part, organizations have offerings in just one of three broad markets:

- Educational assessment,
- Employment testing/screening, or
- Professional certification and licensure testing.

A few companies, however, span two or all three of these markets. The nature of Internet services offered included the following broad categories:

- Internet administration of proprietary tests;
- Internet delivery of the customers' own tests;
- Internet hosting of the customers' own test delivery software; and
- Authoring and/or delivery system software.

The remainder of this section will be organized according to the nature of Internet services companies offer. Within each type of service, offerings to the educational, employment assessment, and certification / licensure markets will be differentiated when appropriate. The list of companies identified below is not exhaustive. Where possible, we have been selective, choosing to highlight those companies whose offerings were considered the most credible or relative.

Internet Administration of Proprietary Tests

Several companies offer delivery of their own or others' proprietary tests over the Internet. Typically, these offerings include test administration, test scoring, and test score interpretation materials (reports). Under this business model, the service provider both 1) operates the Internet test administration system (including the host servers, software, and communications equipment to the point of connection to the Internet), and 2) owns the intellectual property represented in the content of the tests. These kinds of companies are summarized below according to the markets served.

Educational tests

A number of companies offer their own educational tests, primarily for use in the K-12 (kindergarten to grade 12) market. We have chosen to highlight the following companies for one of two reasons. They either 1) represent that their Internet-delivered educational assessments are computer adaptive tests, or 2) have a significant presence in American schools. These companies are mentioned here, not because these tests would necessarily be of use in the STP, but as these companies already have Internet test delivery software that has apparently proven acceptable for use in the schools. Such software could conceivably prove to be adaptable for delivering ASVAB tests.

EdVision.com, based in San Diego, is an Internet publisher of educational materials and services, including a reading placement test that is said to be computer adaptive. The latter is delivered through an Internet site named ReadingPlacement.com to subscribing schools using the schools' own computers and standard Internet browsers. According to the company's description, the reading placement tests can provide teachers or school administrators with a quick assessment of a student's functional reading level. This information is useful for placing students in reading instruction at grade level or in remedial or enrichment reading programs, depending on measured reading proficiency. EdVision.com, while new to the Internet testing arena, is an established publisher of educational materials and databases. It formerly did business as Tudor Publishing. The leadership of EdVision.com includes former senior executives and technical representatives of major educational test publishers, including Harcourt Educational Measurement and Riverside Publishing.

EduTest.com, based in Richmond, Virginia, was formed solely for the purpose of delivering educational measurement instruments and services to schools via the Internet. They have developed both an Internet assessment delivery system, and educational measurement instruments specifically designed for Internet delivery. The principal application of their tests is preparation of school children for tests used in some states for school accountability decisions. EduTest was founded by Dr. Susan Hardwicke, who performed support work during the research and development of CAT-ASVAB in the early 1980s. That early work made her familiar with the characteristics of the ASVAB tests and the special requirements of delivering ASVAB adaptively by computer. EduTest.com was acquired in the spring of 2000 by Lightspan, Inc. Lightspan, founded by the previous CEO of Jostens Learning Corporation, is credible by virtue of the list of major organizations that have invested in it, including Microsoft.

ZapMe! Corporation is a small, public company (formerly Satellite Online Solutions) that provides free satellite-based Internet service to schools and installs the PCs to use it at no cost. ZapMe! runs banner ads and monitors students' web surfing habits for data to sell to advertisers. We mention ZapMe! not because it offers Internet-based testing (it doesn't), but because its stated goals include installing standardized computer labs in 20,000 to 25,000 schools within five years. (Currently, it has installations in 600 U.S. school districts, ostensibly with agreements from additional districts representing thousands of schools and millions of students.)

Each lab contains a local network of 15 Pentium-based PCs, with access to the Internet via a proprietary broadband satellite network named *r)Star*. In 1999, ZapMe! entered into a strategic partnership with Sylvan Learning Systems whereby Sylvan planned to launch tutoring, testing, and other educational services in schools after hours, using the ZapMe! networks. We note that a national network of computer labs, as ZapMe! provides, could provide a delivery vehicle for an Internet version of ASVAB for use in the STP. Encouraging as this is, its future is not clear, as ZapMe! recently altered its business plans, and is in the process of being acquired by an Israeli firm. Likewise, since 1999, Sylvan Learning Systems divested its computer-based testing business to a Canadian firm, Thomson Learning.

Employment tests

There are a number of "dot.com" companies that are offering Internet-based employment testing services. As these companies keep confidential the list of their corporate clients, the extent of Internet-delivered employment testing is unknown. Nonetheless, we have identified several credible employment test service providers. Summaries of these companies and their offerings follow.

e-Selex.com offers employers off-the-shelf, as well as customized systems, featuring Internet delivery, for administering job applications and employment tests, and for evaluating applicants' suitability and likelihood of success in a broad range of jobs. e-Selex offerings include a system for Internet administration of proprietary biographical data (biodata) instruments tailored to specific job categories, as well as custom development of job samples, and skills and ability measures. Although the e-Selex system appears to span the gamut from job analysis to broad-based employment testing to criterion measurement, its core offering appears to be job-specific biodata instruments embedded in Internet-administered employment application blanks. e-Selex offers a broad range of delivery systems , including the Internet, local and wide-area networks (LANs and WANs), as well as administration using stand-alone PCs, fax, e-mail, and print. San Diego-based e-Selex was founded by an industrial psychologist with extensive corporate consulting experience, and special expertise in biodata applications.

Brainbench.com holds itself out as the "leading provider of Internet-based applicant testing services." It claims to have administered over 400,000 assessments in a single year, to applicants to such prominent companies as Computer Science Corporation, Electronic Data Systems, Ernst & Young, and J. P. Morgan. Brainbench appears to specialize in technical tests, oriented toward the information technology (IT) field, although it also advertises knowledge tests specific to a variety of industries, financial and accounting tests, business skills tests, as well as measures of communications ability in five languages, reading comprehension, and English grammar; and a more traditional psychometric assessment of aptitude for programmer / analyst positions. One distinguishing feature of Brainbench assessments is that it uses computer adaptive testing (CAT). According to its promotional materials, its CAT system employs the stratified adaptive ("stradaptive") method pioneered by David Weiss. It is not clear whether its CAT approach is based on item response theory. Brainbench.com is based in Chantilly, Virginia. It was founded by IT professionals.

Reid Systems, recently bought-out by NCS, is an established developer and publisher of paper and pencil integrity tests. It offers its clients access to its applicant assessment instruments via the Internet, accessible at *ReidSystems.com*. The system integrates the application process with non-cognitive testing to predict counterproductive employee behavior such as theft, drug use, absenteeism, and turnover. Reid Systems delivers its products and services through a variety of media, including the Internet, interactive kiosks, automated telephone interviews, stand-alone PCs, fax, and traditional paper-and-pencil. Reid integrity tests are used by a wide variety of companies. Changes in Reid's operations should be followed, since

its merge with NCS could have significant effects upon its business model.

ePredix.com is essentially an Internet-based employment service. Its screening service includes administering employment tests via the Internet to applicants for specific positions. The choice of which employment tests to administer is tailored to the position, on the basis of job analysis data. The tests available apparently include measures of cognitive abilities, personality traits, interests and values, among others. ePredix was founded by an Australian industrial psychologist with extensive human resources management experience. The company is based in San Francisco. At this writing, ePredix is quite new. Portions of the web site are still under development, and online applicant testing is not operational. It is made more credible by a blue-ribbon group of technical advisers that includes some leading academic researchers from the fields of psychometrics and industrial psychology.

However, its chief claim to credibility is based on the origins of its employment tests, many or all of which are established instruments with well-documented psychometric properties. These instruments were acquired from several respected organizations, including Personnel Decisions International, Richardson, Bellows & Henry, and the Vocational Psychology Research service of the University of Minnesota.

DDI, a subsidiary of ASI, Inc., has developed a web-based pre-employment screening device that asks the potential employee a number of questions to determine if the applicant is a good fit for a job (DDI, 2000). Both Bell South and CSX Transportation are using a web-based assessment developed by ASI, Inc. to help in the selection process (ASI, Inc., 2000).

Flex Training is a program under Online Development, Inc. As mentioned before, web-based assessments have become popular to support training initiatives. Many organizations use these types of assessments to help reinforce training and determine the success of training, especially when they are administered in a web-based format. An example of a company that develops web-based assessments as part of a web-based training program is Flex Training. Flex Training develops training-based assessments that allow the customer to determine the percentage of passing grades possible, the number of attempts each individual may have, and whether correct answers will be available to the trainee at the completion of the test (Flex Training, 2000). Once these parameters are set, the assessment system automatically manages itself to the customer's specifications and generates progress reports upon demand.

Internet Delivery of the Customers' Own Tests

To this point, we have discussed only instances in which the Internet testing service-provider administers its own tests, essentially on a subscription basis. In this section, the focus is on tests administered by a service provider using its own software, but which are the intellectual property of the

client. Such tests are available in all four categories:

- Educational tests;
- Admissions tests;
- Employment tests; and
- Professional certification tests.

Although there are not necessarily instances in which all four categories of tests are currently Internet-delivered, each category is discussed below.

Educational tests

Achievement Data, Inc. The Bloomington (Minnesota) public schools, in partnership with Achievement Data, Inc., have developed the Computerized Achievement Level Tests (CALT) that assess children's reading and math levels. They are administered over the Internet to multiple schools in the system (Bloomington Public Schools, 2000). The test is described as a "level test." While not fully adaptive, a level test draws items from the item pool depending on an individual's responses to multiple questions at a similar level of difficulty, rather than changing the level of testing after every question as in a fully-adaptive model. After a determination is made as to which direction the program should branch, a number of questions at that level are presented to the test-taker. One important factor of the test is that a starting point must be determined prior to a student taking the test. A locator assessment is necessary to determine this starting point.

NCS-Vue recently contracted with the state of Minnesota to administer its high school exit examinations in a web-administered format to 8th through 12th graders (NCS-VUE, 2000). Students must pass this exam to graduate from high school in Minnesota. This application is interesting for two reasons. First, it is a high-stakes educational test administered over the Internet. Second, it received significant adverse publicity when NCS-VUE incorrectly scored some tests, erroneously causing some 8,000 students (of 47,000 tested) to receive failing grades (Bakst, 2000). Of these, 54 were high school seniors who would have been denied graduation as a consequence of the error. It is not clear whether Internet administration had anything to do with this mistake, but it is important to recognize that there may be potential dangers in this type of administration.

ASI, Inc., a division of Harcourt Assessment Systems, has developed an assessment for the state of Virginia to determine students' knowledge in chemistry relative to state standards (ASI, Inc. 2000). With the recent widespread implementation of state-mandated, standards-based, testing, Internet delivery of such tests may become common.

Universities and Colleges have been involved in using web-based administration of assessments as well, especially when combined with other web-based initiatives. The most common use for these types of assessments is to assess knowledge relative to coursework. High stakes testing administered in this manner has proven troublesome, causing some instructors and organizations to resort to traditional testing methods for these purposes. Nonetheless, web-based assessments have been found to be useful for developmental purposes, primarily in delivering quizzes to reinforce web-based presentation of content and web-based tutorials (Carbone & Schendzielorz, 1997) Many of these assessments, while

functional, have been primarily experimental in nature and their purpose seems to be one of proving feasibility rather than qualifying as a replacement for other methods of assessment.

Admissions tests

Many academic institutions require applicants to complete tests of knowledge or aptitude as part of the process of selecting applicants for admission. Examples of broadly used computer-administered tests that are delivered by service organizations, but have been developed by client organizations include the Graduate Record Examinations (GRE) and the Graduate Management Admissions Test (GMAT).

Employment tests

To ensure the job relevance of their employment tests, many employers prefer to develop their own tests. While most such tests are administered on the company's premises under their own control, there are some instances of computer administration (but not necessarily Internet administration) by 3rd-party service providers. One example is the Basic Electronics Screening Tool (BEST), a computerized test developed by the Federal Aviation Agency (FAA), administered by Prometric, Inc. in the former Sylvan Testing Centers.

Professional and skill certification tests

Perhaps the broadest use of computer-administered testing is in the area of certification testing, and much of that testing is now conducted over the Internet. For discussion purposes, we can think of certification testing as divided into three categories: Licensure testing, professional certification testing, and technical competency testing.

Licensure testing

In licensure testing, government regulations protect the public by requiring satisfactory test completion to qualify for a license to practice in a professional field. One example of licensure testing is the NCLEX (Nurse Certification and Licensure Examination), which in most states is a prerequisite to licensure as a registered nurse. Although currently administered in Prometric STCs, NCLEX will soon transition to administration by NCS/VUE, which specializes in Internet test delivery.

Currently, a number of states are using web-based assessments for licensure testing in one field or another. ASI, Inc. is an industry leader in providing computer-administered state licensure exams. Some prominent areas in which ASI, Inc. delivers web-based state licensure tests are nurses' aides, insurance professionals, real estate agents and appraisers, among other professions. ASI has recently developed a web-administered testing system called the OMEGA system (Online Multimedia Environment for Global Assessment, ASI, Inc., 2000). This system allows the administration of state licensure exams, as well as professional certification over the web, either in one of ASI's secure testing centers, or directly at the client's site (ASI, Inc., 2000).

Professional certification testing

In professional certification testing, satisfactory test completion is required to qualify for certification by a professional association. These organizations often require individuals to pass an exam for credentialing purposes. In many cases, these organizations are active at the state or even national level. Due in part to the wide

geographic dispersion of their members, more and more of these organizations are turning to web-based administration of their certification exams. An example of computer-administered (but not yet web-based) professional certification testing is the NPTE (National Physical Therapist Examination), which is required by the Federation of State Boards of Physical Therapy for certification as a physical therapist.

Net Certification specializes in developing web-based assessments for professional associations. This company provides training and continuing education programs along with certification testing. Their testing vehicle, eNet Exam, allows both proctored and non-proctored testing environments. NetCertification works with Prometric's secure testing centers to deliver to a proctored environment when necessary (NetCertification.com, 2000).

Technical competency testing

In technical competency testing, an organization such as a private company certifies an individual's technical knowledge in a specific domain on the basis of test performance. Examples abound, particularly in the information technology (IT) field, where companies including Microsoft, Novell, Hewlett-Packard, and Cisco Systems have proficiency certification tests specific to a wide variety of equipment, systems, and software applications, such as computer software packages, network administration, systems management, and others. The most widely used certification of this type is in the use of brand-name products**.**

Virtual University Enterprises, a subsidiary of NCS, is prominent in this area (NCS-VUE, 2000). It delivers certification exams for companies such as Microsoft, Novell, Ericsson, and others. Many of these exams offer the test-taker the option of registering anytime over the web from their home or work computer, but require the individual to actually take the test at one of NCS-VUE's secure test sites.

Brainbench.com, mentioned in an earlier section, also offers computer adaptive tests in technical specialty areas. Its offerings are oriented toward the information technology (IT) field, but it also advertises knowledge tests specific to a variety of industries, as well as financial and accounting tests, business skills tests, and language tests.

Authoring and/or delivery system software

Any organization that wants its tests administered via the Internet must arrange for the tests to be assembled into an Internet-compatible format. This includes the design and formatting of test items for Internet delivery, specification of the sequences of events inherent in the test administration process (such as instructions, rules for examinee-test interaction, item sequencing, branching contingencies such as those of adaptive testing), and myriad other details pertaining to display characteristics, and the user interface.

Internet Hosting of the Customers' Own Test Delivery Software

Some Internet test administration companies limit their offerings to supporting their clients' testing software on Internet servers; a service category referred to as hosting. The business model herein differs from the previous ones in that both the test content and the software that control

test administration have been developed by the client – although the test delivery software may be the product of the hosting company's proprietary authoring system. The service provider operates the Internet server and associated software and communications equipment.

Tables 1 and 2, on the following two pages, summarize the Internet-testing companies described in the previous section.

Study Objective B. Identify major issues surrounding the use of the Internet for employment testing, career exploration, or employment screening

Internet delivery of tests has a great deal in common with other approaches to computerized test administration, such as the use of local area networks of PCs to administer CAT-ASVAB in the MEPSs. A useful framework for examining issues entailed in Internet-delivery of ASVAB is a set of criteria that were formally adopted early in the design and development of CAT-ASVAB. Nine criteria were summarized by Martin and Hoshaw (1997) in their chapter, *Policy and Program Management Perspective,* describing issues to be resolved in the development of CAT-ASVAB. They seem as applicable today in the context of Internet-delivery of the ASVAB as they were when they were first stated in 1982.

These criteria are listed in Table 3, along with short summaries of issues relevant to CAT-ASVAB. Issues relevant to each one are laid out in separate sections following.

Table 3. Nine Criteria Adopted for Formal Evaluation of CAT-ASVAB Prototypes

(Martin and Hoshaw, 1997).

Criterion	Issues
Performance	System response time; graphic display; memory
Suitability	Portability; operating environment
Reliability	Frequency of system failure; provisions to prevent loss of data and facilitate restart/recovery
Maintainability	Maintenance staffing and logistics requirements at test administration sites
Ease of Use	Computer sophistication required of test administrators and examinees
Security	Provisions for protection of test questions and examinee data against compromise
Affordability	Cost-effectiveness compared to the status quo
Psychometric Acceptability	Compliance with professional standards, and equivalence to paper-and-pencil ASVAB
Expansion/ Flexibility	Capability of the system to accommodate future new types of tests

Performance

Three broad performance issues were of special concern in the early development of CAT-ASVAB. They actually preceded the development of personal computers (PCs) as we now know them—system response time, graphics display resolution and speed, and random access memory (RAM) limitations. Early specifications for CAT-ASVAB prescribed a maximum delay of two seconds between an examinee's response to a test question and the beginning

of the display of the next question by the computer system.

In the case of graphics items, the specifications allowed a total of four seconds to complete the display. The intent of such specifications was to avoid distractions, such as slow system response, that might contaminate the measurement process and thus invalidate test scores.

All three concerns have been overcome since the introduction of PCs, by the dramatic improvements in processor speed, graphics capabilities, and both memory and mass storage capacity. The possibility of Internet delivery of ASVAB test items raises these concerns anew. Despite the exponential increases in the power and capacities of PCs over recent years, bandwidth limitations and other characteristics of the Internet raise renewed concerns about performance.

The Internet is not in itself managed or regulated. Thus, it contributes a large degree of variability to any system that distributes or receives information across it. The variability of transmission times over the Internet is one factor that can be expected to affect the apparent rate of system response to input by the test-taker. During any computer-based assessment, the test-taker provides information, that information is processed, and the system responds. Among other things, the processing prior to system response may include item scoring, test score updating, selection of the next item, data storage, retrieval, and transmission.

When the Internet meditates any of these processes, delays in response rates may occur, particularly when different elements of the assessment are stored in different places. For example, the testing program that controls the sequencing and display of test items may be resident on the PC, while the database that contains the actual item pool may reside on an off-site server. In such a system, each item (or sets of items) would have to be downloaded via the Internet to the testing program. Response data for each item, or a number of items, may be sent back to the database for scoring, processing, and storage. The lag that is caused by Internet congestion can significantly affect the time it takes for any one of these processes. To complicate matters, many types of assessments are timed. It is clear that if any test that incorporates timing when its psychometric properties are determined, the delay due to Internet transmission variability can affect system response times, which will impact these psychometric properties. [3]

Response time is a greater issue in computer adaptive testing over the Internet. In non-adaptive assessments, the test taker's computer may upload items in sections or the entire test may be uploaded at once. With computer adaptive testing, responses to each

[3] Although at this writing, ASVAB included two speeded tests, plans are to drop both tests from the battery in the near future.

Table 1. Summary of Developers and Deliverers of Proprietary Tests

Report Page #	Company	Testing Domain	Product - Specific Type of Test	Testing Stakes	CAT?	Delivery Medium	Proctoring Options	Remarks
8	Achievement Data, Inc.	Employment	Job Proficiency	Medium-High	No	Internet	Customer proctored	
6, 11	Brainbench .com	Employment	On-line job applications Work and technical skills certifications Language skills certifications Biodata inventories	Medium to Medium-High	Yes *	Internet	Customer proctored Unproctored from home/office Secure test center option	Some assessments are one-parameter computer adaptive Secure test center option is only voluntary and is available through Prometric
7	DDI, Inc.	Employment	Applicant screening	Medium- High	No	Internet	Unproctored	
5	EduTest .com	Educational	Math and reading achievement level tests	Medium	Yes*	Customer-managed Intranet	Customer proctored	*Adaptive is a level test
5	Ed.Vision .com	Educational	Reading Placement.com (Reading placement test)	Medium	Yes*	Internet	Customer proctored	*Adaptive is a level test
7	EPredix .com	Employment	Applicant screening Selection	Medium to High	No	Internet Stand-alone PC	Applicant screening is unproctored Selection testing is customer proctored	
8	Flex Training	Employment	Testing to reinforce training	Low	No	Internet	Unproctored	
7	Reid Systems (NCS)	Employment	Applicant screening Selection	Medium to High	No	Internet Interactive Kiosks Automated Telephone Stand-alone PC	Non-proctored Customer proctored	

Table 2. Summary of Distributors of Customer and 3rd Party-Developed Assessments

Report Page #	Company	Type of Assessment Distributed	Subcategory of Assessments	Examples of Assessments Distributed	Testing Stakes	CAT?	Delivery Medium	Proctoring Options	Remarks
8	Achievement Data, Inc.	Educational	Educational achievement	Bloomington Minnesota public schools achievement level test	Medium	Yes*	Internet	Customer Proctored	*Level test
9	Harcourt Assessment Systems	Educational	High school achievement	Virginia Knowledge of Chemistry Relative to State Standards	High	Yes	Internet Customer-managed Intranet	Customer proctored	[1]The Omega System is not a test, it is a system
		Employment	Omega System	Licensure/ Certification Testing[1]	High		Vendor-managed testing center	Vendor proctored	[2]The Omega system is CAT capable
10	NCS-VUE	Employment	Certification / licensure	Nurse Certification and Licensure Exam (NCLEX)	High	No	Internet Contractor managed secure testing center[1]	Unproctored Customer Proctored Contractor-proctored (at testing center)	[1]NCS maintains a network made up of testing centers contracted through 3rd parties
8		Educational	High school exit	Minnesota High School Exit Exam	High				
10	Net Certification	Employment	Certification	Numerous examples	Medium-High	No	Internet Contractor-managed secure testing center*	Unproctored Contractor proctored	*Done through Prometric
6, 9	Prometric, Inc.	Admissions	College and graduate school admissions	GMAT, GRE, MCAT, SAT	High	Yes	Vendor-managed testing center	Vendor proctored	
		Employment	Employment Screening	BEST (for FAA)	High				

item may need to be sent back to the server one at a time for scoring, processing, and storage. In addition, a new item must be selected, and transmitted by the server contingent on the response to the preceding question.

Depending on the adaptive testing system software design, adaptive assessment may be more dependent on an unrestricted flow of information over the Internet than a traditional assessment. Internet-based adaptive testing may use more resources, such as data transmission and computational processing than conventional testing, and therefore may be more susceptible to system response delays.[4]

System response time delays over the Internet can be expected to be longer when the items of that assessment contain drawings, pictures, graphs, tables, and other visual media, as opposed to text alone. (The same is true of audio, but CAT-ASVAB currently does not contain audio, so that concern is moot.) Graphics and other media files are large and can take a long time to upload over the Internet. Text files generally take significantly less time.

Graphics elements in assessments present another concern independent of system response time—presentation standards. Current CAT-ASVAB software displays all of its items in highly standardized formats. Any deviations from those formats could alter the psychometric properties of the test items, and thus could affect test validity to some degree. This issue applies to item text as well as graphical components, because typical Internet media displays may vary in terms of text font styles and sizes, display size, aspect ratio, background and foreground colors.

A number of other facets that are tightly controlled in the current CAT-ASVAB software may be free to vary in many Internet applications. Deviations between the current CAT-ASVAB displays and those controlled, say, by an Internet browser, could affect the legibility of text as well as the visual perception and discrimination of graphical components of ASVAB test items.

Currently, test delivery administrators are attempting to address the timing issue. The most popular method is to deliver assessments that are highly dependent on timing in a testing center that has a dedicated Intranet. Unfortunately, this also removes the test from the category of Internet-delivered, which is the focus of the current study.

As a reaction to this problem, some test developers are choosing to forgo timing when developing an assessment. One developer suggested that that in its opinion, the benefit-to-cost ratio for timing was low, and that it was better to forgo the timing requirement in Internet-administered assessments.

Another solution is to set up an on-site testing center with a dedicated Internet connection. Some administrators are creating such centers for large clients. This action reduces the time taken to access the test administrator's central server. While this makes the connection to the administrator's server more robust, it does not guarantee exact timing of items due to the inherent variability of the Internet. This

[4] Workarounds to address this problem typically involve downloading some items in advance of the next questions. This trades-off item security for system speed.

variability would conflict with controlling the exact psychometric properties of an assessment.

Since it is likely that the numerous schools participating in the STP program probably have different ISPs and different levels of Internet access, and also since it is probably not feasible for DoD to provide dedicated Internet connections for the purpose of delivering the ASVAB, this option is probably not a practical administration option.

Technical discussion

As described previously, the administration of an Internet-based assessment is dependent upon many links in a chain, most of which are related to computers and computer systems. The hardware, software, telephony, and computer networks, which the assessment depends upon, must all operate in concert in the two-way delivery of data over the Internet. Two related concepts that are of primary importance, especially when considering response time are those of bandwidth and connectivity. Bandwidth is the amount of information that a computer network can send or receive at any one time over the Internet. Connectivity describes the ability of a user on a computer network to gain access to other servers.

Response rates depend heavily on bandwidth and connectivity. In the case of a computer adaptive test, it may be crucial for the computerized network being used to have large bandwidth so that it can handle the frequent transfer of data that is inherent in CAT. In addition, the user must have good connectivity on-site before the advantages of bandwidth can be realized.

Bandwidth. The bandwidth of a system also limits the numbers of users who can access a particular server and its database at any one time. The greater the number of users who will be testing at any one time and accessing the same server, the larger the bandwidth required to successfully test all the users. If a large-scale computer-adaptive assessment (such as CAT-ASVAB), which is data-intensive, is to be administered over the Internet and involves the concurrent testing of many users, the system involved may need a substantial amount of bandwidth.

Even if a large amount of bandwidth is secured for an application, this will not guarantee standardized performance of the system. One limitation in the ability to capitalize on bandwidth is the dependency of the entire system on the user's Internet connection. Each participating site will most likely have a different Internet service provider (ISP), each with different capabilities. In other words, the uploading of information to the Internet will vary from site to site. Further complicating this process is the fact that the performance from site to ISP, and ISP to the Internet will be variable depending on the level of use at each of these levels.

A second limitation in the ability to capitalize on bandwidth is relative to use of the Internet overall. When there are "spikes" or periods of intense Internet traffic, performance will probably suffer regardless of the amount of resources secured. When this is the case, both uploading and downloading of information would be affected.

Connectivity. Any type of computer-administered assessment requires software that will distribute and display test items. In the Internet assessment industry,

there are currently two types of programs that predominate in that function. These are browser-based and application-based programs.

Browser-based programs have traditionally been used for Internet-based testing. Assessment developers have often used browser software developed by third parties (such as Netscape Navigator and Window Explorer) to distribute HTML-coded assessments. An advantage to this strategy is that most computers already have one of these Internet browsers installed, or they can be installed quickly and in most cases for free. HTML programming is generally straightforward and this type of scripting is generally easy to do. This solution is quick, cheap, and easy.

The are disadvantages to browser-based programs. One is security. HTML code can easily be monitored in transition across the Internet unless it is well encrypted. Another disadvantage is speed. Browser-based programs are generally slower as compared to their application-based cousins since larger amounts of information are going back and forth between the test-taker's computer and the server. HTML code is not as efficient for data transmission as other types of code may be and each page of HTML script must be built individually, unlike the coding for application-based programs. In addition, when Internet browsers are used to distribute assessments, peripheral functions of the browser must be controlled to hinder cheating.

Application-based programs, also often referred to as "Java-based" due to the computer language that is often used to code these programs, use a different method to display the test items to the test-taker. In reality, these programs can be coded using one of the modern object-oriented programming languages such as Java, ASP, or Cold Fusion (Interview with K. Byrne, 8/3/2000). In this scenario, an item-viewing program is loaded onto the test-taker's computer prior to the testing session. The advantage of these programs is that they do not allow viewing of code (Byrne, 2000). Thus, with application-based programs, security is less of an issue. The code that is sent to the user is not easy to intercept and interpret outside of the assessment program, unlike HTML code, which is used in the browser-based programs. Also, timing appears to be more accurate when using application-based programs (Interview with J. Rosen, 8/7/2000).

Currently, some members of the industry are shifting from browser-based testing to application-based testing, particularly because application-based programs appear to be better suited for assessments that require a high degree of item security. Many contractors are beginning to use application-based programs to become involved in high-stakes assessment.

The use of application-based programs for Internet test distribution is relatively new. While a few contractors currently have application-based programs in place, others are currently developing them and plan to have them up and running in the near future. Prometric, Inc. is one company that currently employs both browser-based and application-based programs.

Suitability

In the early development of CAT-ASVAB, two major concerns were subsumed under *suitability*; both had to do

with hardware rather than software. One concern was that a CAT-ASVAB delivery system had to be suitable for use in the existing enlistment testing facilities: Military Entrance Processing Stations (MEPS) and Mobile Examining Team (MET) sites. Suitability for MEPS use implied that CAT-ASVAB equipment could be installed and operated without need for significant modification of MEPS facilities. Suitability for use in MET sites – typically, borrowed rooms used on a temporary basis for testing – implied the use of portable equipment that could be set up quickly, used for test administration, then dismantled and either removed or stored out of sight until the next scheduled testing session.

In the context of Internet administration of ASVAB, the suitability issue is similar, but there are a few additional wrinkles. One issue to be resolved is whether CAT-ASVAB would be administered in the schools, or at off-campus sites. Internet administration in the schools could take place on computer equipment already present in the schools, or on equipment introduced into the schools specifically to administer CAT-ASVAB. Administration at off-campus locations – for example, at facilities such as those of NCS-VUE or Prometric (formerly Sylvan).

In the case of Internet ASVAB administration in the schools, if existing school equipment is to be used, it must be suitable for the purpose. This implies that the equipment must meet critical standards related to compatibility with the Internet test delivery software, and that there must be enough equipment available to accommodate ASVAB STP test administration loads, over and above the normal use of the equipment.

If equipment were to be introduced into schools specifically for ASVAB STP test administration, it could be permanently placed there, or portable equipment could be brought in for each test administration session. These two alternatives raise the same suitability issues that faced CAT-ASVAB planners 20 years ago: Permanent equipment must be suitable to the site and its existing operating environment, without disrupting normal operations, and without need for significant facility modification. Portable equipment must be capable of being set up on a short-term, temporary basis in existing school facilities, then removed at the completion of test administration. Whether the equipment is permanent or temporary, two critical requirements are 1) the availability of adequate electric power, and 2) adequate telecommunication facilities to support Internet access, such as telephone lines or alternative means of broadband data communications.

Reliability

In the present context, the term "reliability" refers to the computer system, not to the psychometric term. The concern about reliability has to do with hardware and software service interruptions and the integrity of data—particularly item response and test performance data collected in the course of ASVAB administration. System reliability is sometimes measured in terms of "mean time between failures." The impact of a system failure may be mitigated by restart or recovery provisions. Data integrity may be protected by data backup provisions, including provisions for redundant storage of test-related data.

It is an axiom of system reliability analysis that the probability of a failure

increases with the number of system components that could fail. In the case of Internet delivery of the ASVAB, the delivery system would potentially have substantially more components than the current CAT-ASVAB system, which for failure analysis purposes consists essentially of a local area network of computers. Failure of any one computer in the network can be mitigated by moving the test-taker to another workstation on the same network, or to a computer operating in stand-alone mode. If a workstation is available, a test can be moved from a failed workstation to the free one in just a few minutes. Redundant data storage provisions in the system virtually guarantee that any loss of test administration data will be limited to the test currently being administered. Data from completed tests are available on the network drive and on removable media, so loss of data is rare.

An Internet-based ASVAB delivery system, however, would be subject not only to failures of the workstation and the local network, but also of the local ISP, the remote server, the data communications network, and devices (such as modems, routers, switches, and data lines).

Depending on the locus of test data storage, and provisions for redundancy, 1) restart / recovery of an interrupted test may be less convenient than that designed into the CAT-ASVAB system; and 2) loss of data from a partially completed test may be more of a risk.

The key point in this discussion is that the topology of an Internet-based system for administering ASVAB is likely to be quite different from that of the current CAT-ASVAB system. For that reason, system reliability is an issue, and one likely to require careful attention if the high degree reliability inherent in the CAT-ASVAB system is required in the Internet-based system.

Maintainability

Maintainability is a counterpart to system reliability, referring to the ability to keep the hardware and software system operating, to fix it when it fails, and to upgrade it from time-to-time as needed. The current CAT-ASVAB system, which has been used to administer ASVAB in all of the MEPSs for several years, has proven to be highly maintainable.

A factor is that all CAT-ASVAB equipment is highly standardized, and the software is centrally maintained. Additionally, the agency responsibility for system maintenance is clear.

ASVAB STP administration over the Internet could prove different, depending on whose equipment is used for test administration, and whose software systems are used for delivery. For example, if schools' computers are used for ASVAB administration, hardware maintenance presumably would be the responsibility of the schools.

While this relieves DoD from the burden of maintenance, it also means that equipment maintenance is outside DoD's control, which could have a negative impact on the STP program if maintenance were deferred or below DoD standards. As another example, if an Internet assessment vendor hosts the ASVAB STP application, maintenance would be that vendor's responsibility, and—although the vendor would have a vested interest in good maintenance—may likewise be outside of DoD control. Finally, various aspects of the Internet itself—routers, switches,

servers, communications lines—may require maintenance from time-to-time. Again, DoD would have no control over such maintenance.

Ease of use

Usability is the quality of a system to be easy to use and understand. In other words, is the system user-friendly? Usability should be considered when developing a Internet-administered assessment system. The system should be easy to use for a number of stakeholders in the assessment (Martin & Hoshaw, 1997), including test administrators, system operators, and the test-takers themselves.

It should be easy to use and understand for test administrators or proctors, especially when they are expected to sign-in test-takers on the system, answer questions, and trouble-shoot minor technical problems. There should be ease of use for those who access the reporting functions of the assessment. Not only do the reports need to be easy to retrieve, but also they should be presented in a manner that is easy to interpret. Additionally, maintenance and trouble-shooting of the system should be straightforward and simple. Finally, for test-takers, the system should be sufficiently easy to use that differences in results between paper-and-pencil and computer-administered testing are negligible.

All of these desiderata have been met in the case of the CAT-ASVAB system, which is currently used only for the Enlistment Testing Program (ETP) in the MEPSs. Some new usability issues can be expected to arise in the event that Internet delivery of the ASVAB is introduced into the STP.

From the test-taker's perspective, there are a number of potential differences between CAT-ASVAB and an Internet-based ASVAB. These include such matters as the user interface, online test instructions, test item display, and the mechanism for responding to test items. Internet use has become so widespread in recent years that none of these is expected to present a major obstacle. Nonetheless, an Internet version of the ASVAB is likely to differ from CAT-ASVAB in a number of its surface features (such as the use of a mouse for input, rather than the special keyboard used by CAT-ASVAB) and these differences will require planning, interface design, and some research.

The impact of Internet-delivery on ASVAB STP test administrators and proctors is likely to be more profound. For one thing, the medium of STP administration would shift from large-group paper-and-pencil testing to (presumably) small group or even individual computer-administered testing. This would not only change the role of the people involved in test administration, it would also change STP test administration scheduling dramatically.

Instead of mass testing in a single 3- or 4-hour time period, the administration schedule of the STP tests in a school would inevitably have to be spread over a longer period—days or weeks—due to limitations in the number of computers available for test administration. While this system may limit the number of individuals who may be tested concurrently, the advantage is to allow continuous testing and/or testing "on-demand" rather than requiring scheduled group examination sessions. In addition, test administrators and proctors would have to both use and troubleshoot the computer equipment and software systems involved

in ASVAB administration, and would need to be aware of and responsive to issues that are specific to computerized, Internet test administration.

Security

Security is one of the most important issues when considering Internet-delivered assessments. The first line of defense against a variety of threats to test validity is the use of proctors to supervise test administration. Proctors can not only guard against cheating; they can also check ID cards, collect test-takers' signatures, and even take fingerprints. However, the ratio of proctors to test-takers may be small if the number of computers used for testing is small. The effect of this is to increase the cost per test.

Other measures used for computerized test administration include storing a digital photograph of the test-taker in the computerized test record, and recording the entire session in a video medium.[5] Emerging technologies will also allow identification by such means as digital scanning of fingerprints, retinal scans, and voiceprints. All of these are technically feasible today, but are relatively expensive. More to the point, they require computers to be equipped with devices not typically found in high school computer labs.

Verification Testing. Verification retesting at the MEPS for STP applicants might be necessary on some or all ASVAB tests. For example, students might be retested on the Armed Forces Qualification Test (AFQT) subtests at the MEPS and then required to retake the entire battery if their retest AFQT score was significantly lower than their score from STP testing. If the incidence of invalidated STP scores turns out to be very low, AFQT testing could be reduced to a sample of STP applicants for continued monitoring. Since at least some of the scores might count, standardized proctoring would be required to check the identity of the examinees and ensure standard administration conditions. A distinct ASVAB form would be used to reduce problems of item compromise associated with on-demand administration of the STP ASVAB.

With respect to the ASVAB, at least three different aspects of security must be considered: Cheating, identity fraud, and item disclosure.

Cheating. Cheating can take many forms. Two forms of particular concern with computer-administered tests are bringing materials/surrogates into the testing site to help take the test, and the use of other functions of the computer to acquire information. One of the most effective guards against cheating is the age-old practice of proctoring. Proctors can ensure that information is not shared or passed from test-taker to test-taker, or between an off-site individual and a test-taker who is on-site (Interview with J. Rosen).

[5] PKI or "public key infrastructure" is an encryption protocol that is often used over the Internet for applications such as digital signatures and certification authorities. This protocol generates a pair of "keys", or passwords, that are contained in the code of the transmissions. Initializing these keys allow an algorithm to encrypt and decrypt data at both ends of the transmission (sender and user). In the case of digital signatures, PKI can authenticate the confidentiality of transmissions between parties, the authenticity of the signer, and of the message, as well as when the signature was activated. An example of dignital signatures is the electronic signature filed with on-line income tax returns.

A new potential form of cheating arises in the context of computerized testing. The test-taker may be able to use the resources of the computer to look up information that would give an unfair advantage. For example, many computers have encyclopedias, dictionaries, and other reference programs residing on their hard drive.

Especially in a Microsoft Windows environment, without proper precautions, a test-taker could easily engage one of these programs concurrently while taking the test to look up information to answer test items. In addition, since the computer is necessarily Internet-capable to administer the assessment, the individual could access many resources available over the Internet to unfairly answer test items (Perception, Inc., 2000).

The most effective way to guard against this type of cheating is to use a testing program that disables many of the other functions of the computer, except for the ones necessary for testing, so that neither other programs on the computer nor the Internet could be accessed. Companies that employ an Internet browser to display items to the test-taker generally use a custom-tailored Internet browser that has limited capabilities as compared to its full-functioning cousins. These "secure browsers" can lockout standard Internet browser functions that allow the test-taker to go forward or backward (when the requirements of the test deem that individuals should not flip back and forth between questions, such as in a computer-adaptive test). The browser can also ensure that no other database is concurrently accessed over the Internet from that browser except for the one running the assessment. This ensures that only the information that is supposed to be displayed during assessment is displayed (Perception, Inc., 2000).

Application-based ("Java-based") programs that do not employ an Internet browser can also be adapted for this purpose. The application-based program can disable all other applications residing on the testing computer, except for the ones necessary for the assessment (Interview with J. Rosen). The application-based assessment program can virtually "take-over control" of the test-taker's computer for the duration of the assessment. This ensures that other programs and other Internet applications and databases cannot be accessed during the assessment.

Access to the testing system itself also needs to be protected to both guard against cheating and to protect test items. Unauthorized individuals must not be able to access the system. The most straight-forward solution to this problem is to use password protection for system access (Interview with B. Tudor). It is also important that passwords be well managed. Prometric, Inc., for example, uses a program called "Proctor Logic." This program monitors a number of registered proctors at each customer's site who are authorized to set-up and monitor assess-ments through their system. Both the proctor's and the test-taker's passwords are verified before a testing session can begin. (Interview with J. Rosen).

Identity fraud. Although the principal use of ASVAB STP scores is for career exploration and counseling, in many instances ASVAB STP scores can be used by individuals to qualify for enlistment in the Armed Services (high stakes). For this reason, it is important to ensure that test-takers are

who they represent themselves to be. Although this is no less an issue in connection with paper-and-pencil testing, administration of the ASVAB over the Internet could present opportunities for identity fraud that are not available in the tightly monitored environment in which the printed ASVAB STP is given. For example, unless safeguards are put in place, Internet-administered tests could be taken unsupervised, with no proctor to verify the test-taker's identity (or to prevent other forms of cheating, such as outside assistance.) Measures to thwart identity fraud are relatively easy to implement, but may be somewhat expensive.

Item disclosure. As the stakes increase in testing, the importance of test item security increases correspondingly. Additionally, the greater the investment of resources used to develop the assessment, the greater the need to protect the integrity of the test. For a test such as the current CAT-ASVAB, which is a high-stakes test where significant resources were committed to develop the assessment, test-item security is at a premium.

Protecting item security on a web-administered test is a complicated process. There are as many potential weaknesses in the system as there are links in the chain that deliver the assessment. Items must be protected at every link in the chain, from where the items reside on the server to protecting against item theft by the test-taker (or others) during administration. In between, tactics must be taken to ensure that items are not stolen while in transit between these two points.

One of the biggest test item security issues related to Internet-based testing is, "where are the test items stored during the actual assessment." In some cases, entire sets of items may be downloaded onto the computer's hard drive. This is not necessarily a significant problem in a low stakes test. On the other hand, this is certainly not a wise practice for a high-stakes test. Most test developers and contractors who develop high-stakes Internet-administered tests load test items only into random access memory (RAM). In this case, the storage of these items only exists so long as the computer is on and the test is being taken. Once the test is completed, or the computer is turned off, the items are erased from the test-taker's computer. This practice hinders theft of items from the test taker's computer.

Information may also be stolen while in transit. The Internet is basically a public domain, and all information transmitted across it is subject to monitoring. It is possible that a computer user with relatively moderate skills could intercept information during the delivery phase. To guard against this, most sophisticated test developers employ encryption methods. Double encryption, currently being used by on-line banks and companies that allow on-line trading of financial products, is used by a number of companies engaged in the development and delivery of Internet - administered assessments.

Information is also vulnerable to theft at the point of the connection of the testing site to the Internet. One solution to this is to have a secure test site that is monitored, while also having a secure connection that impedes hacking into the data stream at that end.

Affordability

At the time of CAT-ASVAB development, "affordability" was a crucial issue in the decision regarding operational implementation. It was in effect defined in cost-benefit terms. The decision to implement CAT-ASVAB hinged on an economic analysis that showed that the costs of implementation (largely equipment costs) were more than offset by resulting cost savings in other aspects of the recruiting and enlistment processing system (Wise, Curran, & McBride, 1997).

Presumably, the prospect of implementing the ASVAB STP over the Internet would be subject to similar economic considerations. A cost-benefit analysis is beyond the scope of this report; however, some aspects of the cost side of the equation can be examined here, albeit briefly.

Components of the cost of Internet administration of ASVAB in the STP include the following: Internet service, web server hosting, testing workstations (computers), test administration costs, software conversion, and research and development. Some of these costs could be nil, if existing resources were used. For example, if schools' computers and Internet connection facilities were used in ASVAB STP administration, there might be no capital costs for delivery system hardware or Internet services. The same would be true if commercial computerized testing vendors' facilities were used for test administration. Test administration costs, on the other hand, might increase substantially over the costs of group-administered printed testing, due either to increased proctoring costs or to per capita costs if commercial vendors were used. There would be costs for web server hosting. These might be direct costs, or might be absorbed into per capita costs if commercial vendors administer the tests.

Many of the costs discussed to this point might be offset to some degree by the elimination of costs related to the printed ASVAB—printing, test material storage and distribution, shipping and processing of answer documents, etc. The costs of proctors might be reduced on the surface, but in fact could be considerably higher. Take, for example, the situation that is likely to occur with widely available Internet testing – several schools would want to test on the same day or about the same time, and conflict with the schedule of OPM testers or other "official" testing personnel. More test administrators or proctors would certainly be needed for CAT-ASVAB over the Internet in the schools.

Additionally, there would be some "one-time" costs, such as costs involved in conversion of the ASVAB STP test materials (instructions and test items) to a digital medium suitable for Internet presentation. Finally, some investment of R&D resources will be needed to address the issue of test equivalence, as well as other psychometric issues that may arise.

Psychometric acceptability

CAT-ASVAB, through its initial and continuing psychometric research and development, has already proven its acceptability with respect to professional testing standards. An Internet-delivered version of ASVAB should have only one hurdle to take advantage of the established psychometric properties of CAT-ASVAB—that of psychometric equivalence. Just as CAT-ASVAB equivalence to printed ASVAB forms had to be demonstrated, the

equivalence of an Internet version to previous versions (whether printed or CAT-ASVAB) would need to be proven, particularly if there are substantial differences between the current CAT-ASVAB and an Internet version. Such differences seem likely, given the expected differences in terms of display characteristics, fonts, graphics, user interfaces, and system performance characteristics.

Expansion/Flexibility

Prior to its operational implementation, the CAT-ASVAB system was required to demonstrate the capability to be expanded to administer non-traditional tests in addition to the established ASVAB tests. This capability has been established, and CAT-ASVAB now includes one such test, a measure of spatial ability called Assembling Objects.

Future expansion of CAT-ASVAB is somewhat constrained, however, by both equipment and operating system limitations. Such limitations need not apply to an Internet version of ASVAB, provided the Internet version takes full advantage of features that are not supported in the CAT-ASVAB system—features such as the capability for sound, enhanced graphics, animation, and streaming audio and video, among others. In short, current Internet capabilities far outstrip the MS-DOS based CAT-ASVAB system with respect to expansion and flexibility.

New Issues

In addition to issues associated with the established CAT-ASVAB evaluation criteria, there are new issues that would have to be considered. A number of these are discussed below.

- Where would the tests be administered?
- When?
- Who would administer them?
- Who would operate the system?
- Test administration agency?

One important question is "Where would the test-takers take the test?" The answer to this question depends on the needs of the organization for which the test is developed, the need for test security, and the desire of the organization for accessibility to the test. Many testing contractors could offer a number of options for the user (organization) in terms of testing site. The user could choose to test the individual in its secure company-managed testing centers, the vendor might offer to set up a secure site at the organization's office, or the user could opt that the test be distributed and taken at an individual employee's workstation or home computer.

Obviously, the degree of security (which includes whether the test is high- or low-stakes) would be one of the deciding factors in the decision. A high-stakes test should not be distributed over an insecure testing site (an employee workstation or home computer). At the same time, it is probably overkill to require a relatively low-stakes test (such as a developmental quiz for skill training) to be administered at a highly secure site such as a contractor-run testing center. Cost, convenience, and accessibility to technical support should also be considered when choosing where to test.

Vendor-managed secure testing centers have the advantage of high security and full-time technical staff who are familiar

with the software and hardware and may trouble-shoot during testing should technical problems arise. Harcourt Assessment, Inc., Prometric, Inc., NCS-VUE, and ASI, Inc., all provide this option to customers. The main disadvantage of this system is that test-takers must travel to an authorized testing center for assessment. This takes away from the "portability" advantage of Internet-based testing.

The second secure option—that the vendor would set-up and possibly manage a secure site at the customer's location—offers the advantage of security and convenience of location. The main disadvantage is that this alternative is more expensive for the customer. It likely would also cause a problem in testing in any large volume.

A special type of secure site that is sometimes offered by a contracting organization is the application of a secure site at a convention or special event such as a job fair. This is a relatively novel approach, and while it appears to offer many advantages in special situations, not enough is known about the practice to make an evaluation.

A customer may also choose to have self-monitored or self-proctored sites. In this case, the customer would set up and be responsible for security at the level of the testing site, while the contracting company would only be responsible for delivering the test (via the Internet and their network) to the site. This is feasible as long as the customer trains and manages the site well. Prometric, Inc. will actually help manage an "authorized proctoring" network. It will help to train and oversee proctors provided by the customer at the site. Achievement Data, Inc. promotes self-managed testing sites (Interview with J. Rosen).

Finally, customers may choose the option of having the test distributed to non-secure sites such as at an employee's workstation or home computer. This situation invites security problems, but maximizes the portability of the assessment. All of the vendors surveyed reported that they provided this option to customers.

Some vendors choose not to involve themselves directly in the responsibility of site security. For example, Ed.Vision.com suggested that customer-monitoring and password protection could ensure the security of an assessment (interview with B. Tudor). It is clear that potential customers need to weigh the amount of involvement that they themselves want in maintaining security, and the costs and benefits of choosing a particular test site.

Issues Relating To Reporting And Interpreting Results

The results of any assessment are the single most important factor to the organization contracting the assessment. Without reliable, accessible, interpretable results, the entire assessment process is virtually useless to the organization.

One of the selling points of computer-administered assessments (which include Internet-administered assessments) is that results can be reported almost immediately after the test-taker is through. Internet-administered testing also provides the attraction of allowing anyone with the proper authorization and an Internet connection to access results instantaneously from multiple testing sites that may be separated by significant distances. Test results do not need to be mailed.

Another advantage of Internet-administered testing is the possibility that all testing data may be immediately processed through one server. All testing data can reside in one place and may be accessed anywhere at anytime. In the past, mainframes at a company's headquarters may have eventually contained complete assessment records, but only after a delay due to regular but intermittent transmission of data to the headquarters mainframe. In Internet-administered testing, the transmission of data is continuous and instantaneous, emanating from the actual test-taker and not a divisional headquarters or a system local to the test-taker.

EdVision.com, Inc. has a novel solution to reporting. It primarily develops web-based assessments for school systems, so reporting is generally directed towards school administrators and principals. EdVision.com sets up a secure physical IP address so that these stakeholders can view data online (Interview with B. Tudor). The system schedules when reports can be viewed and can flag when individuals request reports at times not scheduled as well as flagging when unauthorized individuals attempt to access reports (Interview with B. Tudor).

Issues Related To Updating The Assessment

In most cases, assessment procedures must be regularly updated to account for changes in the criterion being measured, addition of new test items to the item pool, or to implement improvements in the test itself (Martin & Hoshaw, 1997). The issue at hand is the ability of the assessment system to be easily updated. Updating is clearly a lot easier when all you have is a host-server and a central program to change. Items can be conveniently moved in and out of the operational assssment to get data on new items or check-out old, modified items.

The fact that many functions of the assessment system may be contracted out can affect the updating issue. Some companies allow users to have full control over the updating of test items. Users may submit items to the database whenever they like and the changes will be made immediately. Other system management companies act as a mediator in the process. Updated or additional items must be submitted to the company and their programmers and technical staff must make the actual changes. This method produces a lag time between the submission of the change and the actual implementation of it. It is up to the purchasing organization as to the degree of control and the speed of updating that it needs.

A final point is whether the assessment system successfully adapts to changes in it when successive software or hardware upgrades are implemented. This can be driven by the needs of the testing sponsor (an updated version of testing software) or as part of an improvement in the vendor's system (new system hardware or software).

Issues Relating To The Feasibility of a Web-Based Computer Adaptive Assessment

If a computer-adaptive assessment is to be created for web-administration, or if a current computer-adaptive test is to be modified so that it can be administered over the Internet, some important issues must be raised. The first issue is that whatever software will be used must be able to handle the coordination of significant amounts of data, and high rates of transmission, especially when the

assessment is using a 3-parameter item response theory (IRT) model. This puts certain restrictions on which distribution systems can and cannot be used.

Some vendors claim to currently have the capability to run an assessment of this type and distribute it. Achievement Data, Inc., CAT, Inc., ASI, Inc., and EdVision.com, Inc. all are currently using some kind of adaptive testing which they are distributing over the Internet. The Achievement Data and EdVision.com products are currently a very simplified version of adaptive test called a "level test" as described in the previous section. The CAT, Inc. product measures one question at a time, but currently is based on a one-parameter IRT model. When interviewed, these three companies claimed that their current models could be easily adapted for a 3-parameter IRT-based computer adaptive model (such as the CAT-ASVAB).

Other vendors claim that they can modify their system in a relatively short period of time to adjust for computer adaptive assessment of this type. These companies included Achievement Data, Inc., and Prometric, Inc.

Study Objective C. Provide an ordered list of options for prototype development rank ordered from high stakes/high risk scenarios to low stakes, low risk scenarios based on knowledge gained

In this section, we describe recommended options for developing and evaluating three different prototype Internet-STP systems. The three options vary by level of security required for the test scores and the test items. We have not assigned priorities to the options. We are hesitant to do so in the absence of more extensive data on the effectiveness and acceptability of the different options. Recommendations for collecting such data are described in the next section.

The three options we propose are:

Low-security option

This option would not guarantee security for either the scores or the test items. Information provided to recruiters would be similar to results from a screening test; all applicants for enlistment would have to test again on an enlistment form. A distinct, possibly reduced, form would be used for STP career counseling so that exposure of the items would not impact operational enlistment testing. No special procedures would be required to verify the identify of the examinees, since the score would not count for enlistment.

Medium-security option

This option would involve verification retesting at the MEPS for some or all applicants on some or all ASVAB tests. A procedure suggested by Segall may facilitate efficiently verifying unproctored scores at the MEPS (Segall, 2001).

Since at least some of the scores might count, standardized proctoring would be required to check the identity of examinees and ensure standardized conditions. A distinct ASVAB form would be used to

reduce problems of item compromise associated with on-demand administration.

High-security option

This option would continue the present policy of allowing students to enlist on the basis of their STP ASVAB scores. Higher levels of security for both the scores and items would be required.

Criteria for Evaluating Alternative Approaches

We propose two over-arching criteria to be used in evaluating alternative approaches for the Internet-STP against each other, non-Internet computer-based testing approaches, and the current paper-and-pencil approach.

These are:

Impact on recruiting (Primary).

How effective is each option in terms of the number of productive leads generated and any impact on student propensity to enlist relative to the option's costs. Clearly, the major issue in impact on recruiting is the number of qualified leads generated.

Effectiveness for career counseling (Secondary)

To what extent does the option yield valid and reliable career counseling information while minimizing time required of the students and of school personnel and DoD and school costs.

Issues To Be Addressed in Developing Each Option

The development of each of the options will involve different approaches to the key development issues shown in the box on the lest side of the next page.

Low-Security Option

Marketing and distribution

The central notion of this option is to encourage as wide participation as possible, thus generating the greatest number of qualified recruiting leads. Under this option, CD-ROMs containing the basic software would be sent to every school. Education Service Specialists (ESSs) from USMEPCOM would follow up to identify any barriers to use of the new system. The general concept is to eliminate the hassle to downloading software from an on-line site, although a download option should also be created. In either case, updates could be detected and downloaded once the base software is installed, eliminating the need for manual distribution of new versions.

Development Issues

Marketing and distribution

Would software be distributed to schools or could everything required be downloaded from a central web-site? What level of password protection would be required and how would passwords be distributed and monitored?

Test content

Would the complete ASVAB be administered or could some of the tests be shortened or eliminated? DMDC currently plans on administering the entire battery. Note that the STP currently uses only the AFQT subtests to create the ASVAB code used with the career counseling materials.

Test delivery software

This includes the architecture of the software, what programs and data might be permanently resident on local computers, and what would be downloaded as used, what copy-protection could be used for transient information.

Monitoring and proctoring

Who would proctor test administration? To what extent would the provision of feedback to students be monitored?

Adequacy and accuracy of identifying information

What information would students have to provide to DoD to participate? To what extent would this information be verified before the student can proceed with testing?

Score Reporting and Counseling

How would score feedback be provided (immediately or after authentication)? How would score interpretation and career counseling be handled? What printed material would still be required?

Test content

Under this approach, test scores would not be used for enlistment so test content could be reduced to ensure that nearly all students could easily complete the battery in a single class period. We suggest convening a panel of career counseling experts to rank the importance of the different tests and consider the level of accuracy needed for the scores from each of these tests.

Test delivery software

Under this option, both software and encrypted item data could be resident on the local machine for speed of processing. We expect that a system that could run under most versions of Windows would suffice, but it might be desirable to create a separate version for MACs. It is possible that local storage requirements would be a problem. We do not expect this to be the case, but it should be checked during pilot testing.

While testing could run largely off-line, Internet connection would be required to ensure that student information is captured. A connection would be required before testing were allowed to begin to transmit student identifiers. A second connection to upload item response and score information would occur after testing but before score information were presented to the students. Once testing were completed, student information would be saved on diskette or to the local hard-drive for input to a subsequent career exploration session using a modified version of the ASVAB Career Exploration System (ACES).

During development, it would be decided whether to include a browser or to create options for using existing Netscape or Internet Explorer browsers. The approach taken should be designed to allow as many schools to participate as possible.

Monitoring and proctoring

Since scores would not count for enlistment under a low stakes option, proctoring could be eliminated altogether. Obviously the software would have to be as "foolproof" as possible. Additionally, some provision for reporting unexpected problems and getting assistance (perhaps through the ESSs) would be needed. In the past, recruiters have frequently been used as proctors. This provided them with an opportunity to be seen on campus, but there was frequently little chance for meaningful dialog with the students. By freeing recruiters from this responsibility, they might be able to schedule more time for activities with greater opportunity to interact with students.

Adequacy and accuracy of identifying information

A key concern with the low-security approach is that students might not be motivated to provide accurate identifying information. It would be desirable to create an option for mailing follow-up information to students, or distributing it through the schools. Students would only receive the follow-up information if they provided correct identifiers. Options range from more extended score reports to generic information about civilian and/or military career exploration not tied to their scores. During piloting, focus groups and follow-up visits with school officials would be used to determine students willingness to provide accurate information.

Score reporting and counseling

In keeping with the concept of making this option as quick and easy to use as possible, outside help would not be required for score reporting and interpretation. Materials such as the *ASVAB Student Workbook* and *Counselors Guide* would be available to help counselors in supplementing the information presented by the software.

Expected Benefits

The primary benefits expected from the low-security option would be increased student participation and possibly reduced costs. It is expected that savings from reduced personnel time for proctoring and score interpretation would more than offset the costs of having to administer a short verification test at the MEPS for those students who wished to enlist. The chief concern is that this approach might have minimal, if any, impact on students' propensity to enlist and that the need to retest could create an additional barrier. Potential inaccuracy in identifying information or test scores is another possible concern that would have to be carefully monitored.

Medium-Security Option

Marketing and distribution

While software might be distributed in much the same way as under the low-security option, additional procedures would be required to ensure that each testing session was appropriately proctored. Since this option includes verification retesting (See Segall, 2001), we believe it might be possible to use school personnel as monitors. Under such an arrangement, counselors would be required to sign an agreement and receive some training to obtain a password needed to run the software. The agreement would stipulate the conditions under which the test could be administered and the counselor's willingness to warranty the validity of each testing session. ESSs could be used to collect signed agreements and distribute

(counselor-specific) passwords, but more extensive interaction with schools would be required before testing could begin. Education Service Specialists (ESSs) from USMEPCOM would follow up with telephone calls or visits to identify any barriers to use of the new system. In keeping with the medium-security nature of this option, an option might be provided for counselors to enroll on-line and receive their passwords by e-mail. An assessment of legal and regulatory acceptability of the use of non-federal governmental personnel would have to be made.

Test content

Under this approach, a full CAT-ASVAB or verified reduced length version of the ASVAB would be administered, since scores could be used for enlistment.

Test delivery software

Test delivery software could be largely the same as under the low-security option. Two additional requirements would be added. First, some information about the testing platform would be uploaded at the beginning of each session to ensure that minimum requirements were met. Second, counselor input would be required at the beginning of each testing session and again at the conclusion before score information was reported. Another way that security could be increased under this option is through either deeper encryption of the test items or, preferably, through downloading item information during the test session so that items would not be locally resident.

Monitoring and proctoring

Under this option, students would not be free to take the test on their own. Supervision by an authorized person would be required. Special visits by ESSs or other DoD personnel would not be required, however, so that testing could be run on-demand throughout the school year. Some training would be required to ensure that monitors understood their role and responsibilities and could handle most problems that might occur during testing sessions. This training might be a pre-condition for receiving administrator pass-words. Each administrator would be required to obtain a new password each year providing an opportunity for some level of re-training.

Adequacy and accuracy of identifying information

Because an administrator would confirm the information submitted by each student, special incentives (e.g., mail-out items) might not be required under this option. However students would almost surely have an option for requesting that their information not be given to recruiters, so some incentives might still be useful. For example, pop-up calendars could be sent to students who allow DoD have name and address information.

Score reporting and counseling

The medium-security option would likely include a greater role for school counselors in score interpretation. Counselors would be required to confirm test administration validity at the end of teach administration, so they would be physically present to assist with interpretation and use of score information.

Expected benefits

The primary benefits expected from the medium-security option would be increased accuracy in the information provided by the students. Assuming school personnel

involvement, DoD costs would still be minimized since its personnel would not have to be present at testing or interpretation sessions. This option would thus also support year-round on-demand testing enabling higher rates of student participation in comparison to current, fixed-session procedures. As with the low-security option, a potential concern about this approach would be that, without the presence of DoD personnel, there might be little impact on the propensity of students to enlist.

High-Security Option

Marketing and distribution

Under the high-security option, test sessions would continue to be monitored by DoD personnel. Thus, ESSs would continue to work with each school in scheduling testing sessions and arranging for proctoring. Software could be brought to each testing session by DoD personnel and secured by them at the end of each session to prevent unauthorized access to the software.

Test content

Under this approach, as with the medium-security option, a verifiable CAT-ASVAB would be administered, since scores could be used for enlistment.

Test delivery software

Under the high-security option, developers could explore the full range of options for local versus Internet provided software components. This option could involve only minimal use of the Internet, to speed and centralize collection of student information. Test administration and scoring could be handled by stand-alone software, since use of the software would be closely monitored. Nonetheless, developers might choose a higher level of Internet use to check for and download updates at the beginning of each testing session and to centralize storage of item and other data so as to facilitate updates to the item bank. It might also be desirable to use this vehicle to field test new items and centralized storage of item information would facilitate routine rotation of the experimental items.

Monitoring and proctoring

Under the high-security option, all testing sessions would be proctored by DoD personnel. New regulations might be required as to the minimum ratio of proctors to students. With computer administration, proctors would not be required to read instructions nor to start and stop each test and this could change proctoring requirements somewhat. Piloting would be required to determine whether more stringent monitoring would be required under a CAT-ASVAB scenario where students might be working on different tests and finishing at differing times.

Adequacy and accuracy of identifying information

DoD proctors could provide more stringent checks of student identities. In addition, proctors would have an opportunity to explain the benefits of allowing recruiter access to test information, so there would be a reduced need for additional incentives.

Score reporting and counseling

A key feature of computer-based test administration, is that scores are available as soon as the battery is completed. Since DoD personnel would proctor each test session under the high-security option, they would

also be available to aid in score interpretation. Under an adaptive version of ASVAB, however, students would finish at different times. Consequently, it might still prove desirable to schedule separate score interpretation sessions in many cases.

Expected benefits

The primary benefits from the high-security option in comparison to current procedures would be reduced student testing time through adaptive testing and immediate collection of score information. The reduced testing time would make it easy to schedule sessions and would reduce time requirements for proctoring each session.

The primary benefits in comparison to the low- and medium-security options would be eliminating the need for retesting to enlist and providing for a greater on-site presence of DoD personnel resulting in more opportunities to increase enlistment propensity. Cost and some limitation in the number of students who would participate are the greatest concerns with the high-security approach.

Study Objective D. Identify a milestone schedule for each of the options and provide detailed listing of tasks necessary to accomplish the implementation of options selected for the milestone schedule by the monitoring agency

In this section, we outline the milestones that would need to be addressed to ensure a successful implementation of an Internet-based STP. The milestones are in many cases independent of the specific option selected from those we proposed. Regardless of security option chosen, test and software development and implementation should proceed as we describe below. Most of the milestones we propose result from our belief that fielding an Internet-based system available to all high schools nationally requires that we develop and test the software in accordance with standards prescribed for developing military software systems.

Because of the central role that software development would take in fielding the Internet STP, we propose a relatively detailed set of milestones for that one activity, organized around the phases of the acquisition cycle:

- Concept exploration and definition
- Demonstration and validation
- Development
- Production and deployment
- Operations and support

Supplementing the more detailed software-development milestones, we briefly outline milestones for those critical tasks that may be more familiar to the test-development community:

Item pool development and evaluation

download option should also be created. In either case, updates could be detected and downloaded once the base software is installed, eliminating the need for manual distribution of new versions.

- Support materials
- Operational test and evaluation

- Training MEPS and Service personnel
- Training school personnel

Software Development Milestones

The purpose of the first few phases of the software acquisition life cycle would be to ensure that the software underlying the Internet STP was developed in a structured way, and that it would ultimately meet requirements. It is important to recognize that during software development, errors would be made and opportunities would be revealed. Similarly, changes would be proposed, and changes would be changed. Unless carefully controlled, the resulting complexity would ensure that the development process would be error-prone, time consuming, and expensive. For this reason, the concept exploration and development phase might be the most important phase—by exploring and testing alternatives in a systematic way, without focusing too early on "the answer," risk would be reduced in the later phases, where "false starts" would be much more expensive.

Concept exploration and definition phase

This phase of the acquisition life cycle would include developing the information and analyses necessary to identify and evaluate alternative functional and technical concepts that satisfy the mission. It would be during this phase that various approaches were tested on a limited basis for feasibility. Many of the questions raised above should be explored and answered during this phase. Specific milestones within the concept exploration and development phase include the following:

- Identify functional requirements
- Assess alternative functional concepts
- Assess alternative technical concepts
- Select best program concept
- Select program procurement strategy
- Perform risk assessment analyses
- Develop functional description
- Develop demonstration / validation plan

Demonstration and validation phase

This phase of the acquisition life cycle includes designing, coding, testing, and demonstration of functional capability subsets. Specific milestones within this phase include:

- Complete software requirements analysis
- Develop preliminary design
- Develop detailed design
- Code demonstration program
- Test demonstration program

Development phase

During this phase, all user-agreed capabilities are satisfied and the software would be prepared for deployment. The activities in this phase would be repeated for each functional capability subset until the

overall design is satisfied. Specific activities include:

- Code and test
- Integrate and test

The following phases of the software development life cycle would be critical to the successful fielding of an Internet STP. Success or failure of the program—measured by satisfaction of users and their use of the system—would depend as much on activities performed in these phases as it would on earlier phases of development.

Production and deployment phase

This phase consists of producing and deploying the system. Here as well, the system's performance would be evaluated and corrected as necessary. Specific activities include:

- Develop management transition plan
- Develop support plan
- Develop post-deployment operational assessment plan
- Assess overall software development
- Assess overall risk

Operations support phase

This phase consists of activities necessary to operate and maintain the system throughout its life cycle. Specific areas needing to be addressed include:

- Correcting malfunctions
- Assuring security safeguards
- Updating operating procedures
- Reassessing affordability, performance, and benefits
- Deciding on modernization or termination

Test Development and Implementation Milestones

Producing an Internet STP involves far more than software development, of course. All the usual ASVAB development milestones would need to be met as well. In this area, at least, there are well-established procedures within the ASVAB development community that would guide the process. For that reason, we provide only a brief outline of some of the milestones that would take on increased importance as the software development process proceded.

Item pool development and evaluation

The software's fundamental purpose is to display items for the examinee. Those items would need to be developed, or provisions would need to be made to use items that already existed. Issues here would depend heavily on decisions made concerning test security requirements and whether or not the Internet STP perhaps would be a computer adaptive test. For example, a low-security option, coupled with some kind of verification test, would substantially reduce the number of items needed, and might allow simple reuse of retired items from previous operational tests. In any case, current procedures exist to develop items, evaluate their quality, subject them to sensitivity reviews, etc.

Psychometric issues

The usual psychometric issues in ASVAB development would need to be confronted in the Internet STP system. Here such concerns as item calibration and test form equating would have to be resolved. In addition, issues of computer hardware comparability, browser comparability, and so forth—as addressed in earlier sections of this report—would require study and resolution. Experience with the CAT-ASVAB development process should guide the development of milestones for resolving psychometric issues.

Support materials

As with the current STP, materials for schools, students, and recruiters would need to be developed and tested. Experience with producing and maintaining existing STP support materials could guide this activity. Specific milestone schedules depend on answers to the many questions we raise in earlier sections.

Operational test and evaluation

This phase of ASVAB development is well understood when it comes to psychometric issues. The Internet STP adds levels of complexity to the process. We strongly encourage a strategy of limited implementation and testing of the system before widespread distribution takes place. Again, the CAT-ASVAB experience could inform the milestones, with the added complexity of the need to deal with school systems.

Training MEPS and Service personnel

Internet STP would place new demands on MEPS and Service personnel. Under some implementation scenarios, they would spend far less time proctoring tests. At the same time, they might be confronted with technical questions that they would have to be prepared to answer.

Training school personnel

The current STP provides school personnel with print materials and in-person help in interpreting test scores. There have also been many instances of in-person training provided directly to counselors and teachers. Providing an Internet STP would carry with it the requirement to expand the training provided to school personnel. Under virtually any implementation scenario, school personnel would be required to be knowledgeable about the system—at least enough to help students connect with the testing site.

Recommendations/Conclusions

The possibility of conducting DoD Student Testing Program ASVAB testing via the Internet is intuitively appealing. Our survey of existing Internet-based testing programs provides ample examples to support the idea that an Internet STP could be technically feasible. At the same time, it is clear that DoD and the Services—and participating schools and students—would benefit from many advantages that this testing mode would bring.

If the Service recruiting commands were willing to implement a program that did not provide enlistment-eligible test scores for either selection or classification (the low security option we described

earlier), or, the low secuity option with verification testing (essentially a medium security option), there could be large advantages for testing. Dr. Dan Segall of DMDC presented a method for efficient verification at the National Council on Measurement in education (NCME) meeting in Seattle in April, 2001 (Segall, 2001). A tryout using such a procedure would be necessary.

With no need to proctor the exams, the cost of OPM test administrators for the STP might be eliminated. Similarly, the time now necessary for recruiters to proctor STP exams could be returned to the Services. While no dollars flow directly back to the recruiting commands, more recruiter time might become available for much higher priority tasks—actually meeting with students and counselors, following up on leads generated by the program, etc.

The same rationale applies to USMEPCOM and Service ESS—without the need to spend time scheduling schools, Service proctors and ESS could spend time performing tasks that directly contribute to recruiting success—marketing the program to schools, parents, and students; training counselors, teachers, and recruiters in score interpretation; interacting with students at the high school. Printing and distribution costs for test booklets and score reports could also be reduced.

Beyond direct cost reduction, an Internet STP administered under a low-security option could reasonably be expected to directly help recruiters by increasing the number of schools using the program, and increasing the market penetration within schools.

Many schools would find on-demand testing attractive. Without the need to schedule the entire school on one or perhaps two days per year—a task that many schools find impossible to accommodate—more schools might be likely to at least offer program to students.

If an adaptive version of the program were offered, many schools (and students) would find the resulting time savings extremely attractive. This particular advantage of an adaptive STP ASVAB would likely increase participation, regardless of the security option chosen for the program.

Other advantages could accrue to the Services from converting to an Internet STP. These advantages might not even be dependent upon the particular security option chosen for the program. For example, any foreseeable solution for an Internet STP would likely provide faster score turnaround—and thus faster leads—to recruiters. Similarly, the military's desired "high-tech" image would be improved simply by the move to more cutting edge testing systems.

Certainly, there are questions that must be answered before any final determination on implementing an Internet STP. We identify many of the critical questions in this report. Some of these questions might be answered based on policy concerns (for example, the acceptability to recruiting commands of offering an STP where scores could not be used for enlistment).

Other questions could and should be answered only after feasibility studies, small-scale demonstrations, and tryouts. For example, there might be hardware or software constraints imposed by some schools. There might be Internet access concerns with regard to other schools. Some technical decisions might be made

based on download speeds typically available in schools. In each of these areas, program planners would need to gather appropriate information from the schools themselves, recruiters, and potential applicants.

Considering the large number of stakeholders involved with the STP, (e.g., DoD and the Services, high schools and community colleges, USMEPCOM, recruiters, the testing community, and school counselors), DMDC should conduct a full-scale market survey to gather input from potential users of an STP-Internet system, to estimate the impact, and to assess need and demand.

Ideally, a broad set of models would be compared, ultimately leading to a larger-scale evaluation of one or a small number of potential models. This incremental approach to assessing the potential implementation of the STP ASVAB via the Internet would likely take a minimum of two to three years.

There would be up-front costs incurred before any Internet STP could be available. We believe these costs could be recovered by the continuing savings that would result from an Internet STP. Here, too, demonstration projects would provide data necessary to confirm or deny our belief. Some of the costs that could be avoided by an Internet STP might be easily estimated (e.g., OPM examiner costs).

A critical aspect of the overall program, however, would be its recruiting effectiveness, i.e., the number of qualified leads generated. The cost-effectiveness of investing in system development would ultimately hinge on the impact of that system on recruiting success. Without demonstration projects, it would be impossible to accurately assess the program's impacts—both intended and unintended.

Finally, policy-makers would be well-advised to remember that there are costs associated with doing nothing. The STP has many critics. Other testing programs continue to take market share away from the program, despite the STP's obvious cost advantages.

Maintaining a cutting edge testing system is one approach to keeping the overall program viable. The trend today is clearly toward computer-based testing in general. An Internet STP would keep the program at the forefront of testing systems.

The enlistment testing program has proven the advantages of CAT-ASVAB. Demonstration projects turned out to be critical components in generating enthusiasm for the program within USMEPCOM and the Services. An Internet STP would similarly benefit from demonstrations and tryouts.

References

Achievement Data, Inc. (2000). Welcome [on-line website] Available: http://www.achievementdata.com/

ASI, Inc. (August 1, 2000). ASI homepage [on-line website] Available: http://www.asisvcs.com/asihome.

Bakst, B. (August 18, 2000). Diplomas awarded after testing error. Washington Post Online [on-line newspaper] Available: http: washingtonpost.com

Bloomington Public Schools. (July 24, 2000). FAQ about our tests. [on-line website] Available: http://www.bloomington.k12.mn.us/Staff_Resources/Office_of_Research_and_Evalu.../faq.htm

Carbone, A., & Schendzielorz, P. (1997). Developing and integrating a web-based quiz into the curriculum. Proceedings from WebNet 97 World Conference of the WWW, Internet & Intranet Proceedings (available on ERIC, No. ED429531).

DDI. (July 28, 2000). Net-Interview [on-line website] Available: http://www.advantagehiring.com/about/pip/netinterview.html

EdVision.com. (July 27, 2000). Readingplacement.com [on-line website] Available: http://www.readingplacement.com/company.htm

Flex Training. (August 31, 2000). Web-based training solution by Flex Training. [on-line website] Available: http://www.flextraining.com

Hamilton, L.S., Klein, S.P., & Lorie, W. (2000) *Using Web-based testing for large-scale assessment*. Santa Monica, CA: RAND Education.

Martin, C. J., & Hoshaw, C. R. (1997). Policy and program management perspective. In W. A. Sands, B. K. Waters, & J. R. McBride (Eds.), *Computer adaptive testing: From inquiry to operation* (pp. 11-20). Washington, DC: American Psychological Association.

NCS-VUE. (July 31, 2000). Virtual University Enterprises [on-line website] Available: http://www.vue.com/

NetCertification.com, Inc. (2000). NetCertification Distance Learning, Online Training [on-line website] Available: http://www.netcertification.com

Sands, W.A., Waters, B.K., & McBride, J.R., (Eds.) (1997). *Computer adaptive testing: From Inquiry to Operation*. Washington, DC: American Psychological Association.

Segall, D.O.. (April, 2001). *Measuring test compromise in high-stakes computerized adaptive testing: A Bayesian strategy for surrogate test-taker detection*. Paper presented at the annual meeting of the National Council on Measurement in Education. Seattle, WA.

Wise, L.L., Curran, L.T., & McBride, J.R. (1997). CAT-ASVAB cost and benefit analyses. In W. A. Sands, B. K. Waters, & J. R. McBride (Eds.), *Computer adaptive testing: From inquiry to operation* (pp. 227-236). Washington, DC: American Psychological Association.

APPENDIX A

Addresses of Internet Testing Referenced Sources

Provider	Address	e-mail	Phone
Achievement Data, Inc.	Achievement Data, Inc. 555 West 78th Street – Suite E Edina, MN 55439-2702	achievementdata.com	952-946-1854
Harcourt Assessment Systems, Inc. (PsychCorp)	Harcourt Assessment Systems, Inc. Three Bala Plaza West Suite 300 Bala Cynwyd, PA 19004	asisvcs.com	877-374-1153
Brainbench, Inc.	Brainbench, Inc. 14901 Bogle Drive Chantilly, VA 20151	brainbench.com	703-437-4800
Computer Adaptive Technologies	(CAT, Inc.) 1007 Church Street Evanston, IL 60201 USA	catinc.com	800.255.1312
(DDI) (Advantage Hiring, Inc.)	Advantage Hiring, Inc. 1225 Washington Pike Pittsburgh, PA 15017-2838	advantagehiring.com	800-726-7136
EduTest, Inc.	EduTest, Inc. 6800 Paragon Place Suite 237 Richmond, VA 23230	edutest.com	804-673-2253
EdVision.com Corp.	EdVision.com 1320 Columbia St.- 3rd Floor San Diego, CA 92101	edvision.com	800-998-4531
e-Predix, Inc.	301 Mission Street Suite 200 San Francisco, CA 94105	epredix.com	415-615-0200
e-Selex	E-Selex.com 939 Hibert St., Suite 107 San Diego, CA 92131	e-selex.com	858-695-6279
Net Certification	Net Certification 3530 Forest Lane – Suite 113 Dallas, TX 75234	netcertification.com	800-295-1685

Addresses of Internet Testing Referenced Sources
(Continued)

Provider	Address	e-mail	Phone
Online Development (Flex Training)	Flex Training 13555 Automobile Blvd. Suite 350 Clearwater, FL 33762	flextraining.com	----------
Prometric, Inc. (Thompson Learning)	Prometric, Inc. 1000 Lancaster Street Baltimore, MD 21202	prometric.com	866-prometric
Reid Systems (NCS)	Reid Systems Online 158 West Ohio Street Chicago, IL 60610-4210	reidsystems.com	800-922-7343
Virtual University Enterprises NCS Pearson, Inc. (NCS-VUE)	NCS-VUE 1100 Prairie Lakes Drive Suite 300 Eden Prairie, MN 55244-3857	vue.com	952-995-8800
ZapMe! Corp.	ZapMe! Corp. 3000 Executive Parkway Suite 150 San Ramon, CA 94583	zapme.com	877-373-3728