



Calhoun: The NPS Institutional Archive
DSpace Repository

Faculty and Researchers

Faculty and Researchers' Publications

2004

Test Consumers in the Military Use of the Military Career Exploration Programs in Schools

Laurence, Janice H.

<http://hdl.handle.net/10945/40470>

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun



Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

<http://www.nps.edu/library>

*Measuring Up: Assessment Issues for
Teachers, Counselors, and Administrators*

Edited by
Janet E. Wall
Garry R. Walz

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
 - Minor changes have been made to improve reproduction quality.
-
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

45 46

MEASURING UP: Assessment Issues for Teachers, Counselors, and Administrators

Edited by

Janet E. Wall, EdD & Garry R. Walz, PhD, NCC

BEST COPY AVAILABLE

ERIC
CASS

Co-Published by
ERIC Counseling & Student Services
Clearinghouse

ERIC
nbcc

and the National Board for Certified Counselors

Measuring Up

Assessment Issues for Teachers, Counselors, and Administrators

Edited by Janet E. Wall and Garry R. Walz

Copyright © 2004 CAPS Press

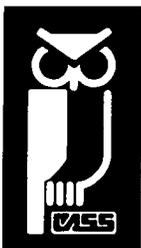
P.O. Box 35077

Greensboro, NC 27425-5077

All rights reserved.

ISBN 1-56109-099-9

This publication is funded in part by the U.S. Department of Education, Office of Educational Research and Improvement, Contract No. ED-99-CO-0014. Opinions expressed in this publication do not necessarily reflect the positions of the U.S. Department of Education, OERI, or ERIC/CASS.



BEST COPY AVAILABLE

Contents

Preface	xi
<i>Janet E. Wall & Garry R. Walz</i>	
Introduction	xxi
<i>Thomas Clawson</i>	
About the Editors	xxiii
About the Contributors	xxv
Abbreviations	xliv
Section A. The Basics of Testing	
Chapter 1	
Why Use Tests and Assessments?	3
<i>John Fremer & Janet Wall</i>	
Chapter 2	
Types and Uses of Tests	21
<i>Timothy Vansickle</i>	
Chapter 3	
Guidelines for Selecting Appropriate Tests	33
<i>Patricia Jo McDivitt & Donna Gibson</i>	
Chapter 4	
Reporting and Interpreting Test Results	53
<i>Deborah J. Harris</i>	
Chapter 5	
Informing Test Takers	65
<i>William D. Schafer</i>	

Section B. Assessment Issues for Special Populations and Audiences

- Chapter 6
Fair and Valid Use of Educational Testing in Grades K-12 81
Janet E. Helms
- Chapter 7
Test and Item Bias
What They Are, What They Aren't, and How to Detect Them 89
Barbara B. Ellis & Nambury S. Raju
- Chapter 8
Racial and Ethnic Difference in Performance..... 99
*Nathan S. Hartman, Michael A. McDaniel
& Deborah L. Whetzel*
- Chapter 9
A Test User's Guide to Serving a Multicultural Community.. 117
David Lundberg & Wyatt Kirk
- Chapter 10
Lost in Translation
Issues in Translating Tests for Non-English Speaking, Limited English Proficient, and Bilingual Students 127
Sharon M. Goldsmith
- Chapter 11
Testing Students With Limited English Proficiency 147
Kurt F. Geisinger
- Chapter 12
Inclusion of Students With Disabilities in State and District Assessments 161
Martha L. Thurlow & Sandra J. Thompson
- Chapter 13
Assessment of and Accountability for Students With Disabilities
Putting Theory Into Practice 177
Judy Elliott

Chapter 14	
Assessing Students With Serious Mental Health and Behavioral Problems	
Clinical Assessment for Educators	197
<i>Jo-Ida C. Hansen & Amy L. Conlon</i>	
Chapter 15	
Broadband and Narrowband Measures of Mental and Behavioral Health	
Counseling Assessment for Educators	213
<i>Amy L. Conlon & Jo-Ida C. Hansen</i>	
Chapter 16	
Assessment of Family Issues	
A Guide for Educators	231
<i>Craig S. Cashwell & Randolph H. Watts, Jr.</i>	
Chapter 17	
SUBSTANCE-Q	
A Practical Clinical Interview for Detecting Alcohol and Other Drug Abuse	245
<i>Gerald A. Juhnke & William Bryce Hagedorn</i>	
Chapter 18	
Overcoming Test Anxiety	
Giving Students the Ability to Show What They Know	257
<i>Brian Goonan</i>	
Chapter 19	
Assessment for Children Ages 3 to 8 Years	273
<i>Marcy Priess Guddemi</i>	
Chapter 20	
Issues in College Admissions Testing	283
<i>Julie P. Noble & Wayne J. Camara</i>	
Chapter 21	
Assessment and College Course Placement	
Matching Students With Appropriate Instruction	297
<i>Julie P. Noble, Jeff L. Schiel, & Richard L. Sawyer</i>	

Chapter 22	
Test Consumers in the Military	
Use of the Military Career Exploration Program in Schools	313
<i>Janice H. Laurence</i>	

Section C. Special Topics and Issues in Assessment

Chapter 23	
Educational Assessment in a Reform Context	325
<i>Michael H. Kean</i>	

Chapter 24	
Education Assessment in an Era of Accountability	335
<i>Peter Behuniak</i>	

Chapter 25	
Applications of Professional Ethics in Educational Assessment	349
<i>Pat Nellor Wickwire</i>	

Chapter 26	
Educational Testing Integrity	
Why Educators and Students Cheat and How to Prevent It	363
<i>Gregory J. Cizek</i>	

Chapter 27	
Practice Tests and Study Guides	
Do They Help? Are They Ethical? What is Ethical Test Preparation Practice?	387
<i>Carole L. Perlman</i>	

Chapter 28	
Test Preparation	
What Makes It Effective?	397
<i>Jeff Rubenstein</i>	

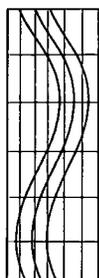
Chapter 29	
Professional Testing Standards	
What Educators Need to Know	417
<i>Wayne J. Camara</i>	

Chapter 30	
Training Educators to Develop Good Educational Tests	427
<i>Patricia Jo McDivitt</i>	
Chapter 31	
Assessment Competencies for School Counselors	443
<i>Patricia B. Elmore & Ruth B. Ekstrom</i>	
Chapter 32	
Test User Qualifications	
Who Can Use What Tests?	453
<i>Thomas Warren Clawson & Wendi K. Schweiger</i>	
Chapter 33	
Assessment for Learning	
Classroom Assessment to Improve Student Achievement and	
Well-Being	463
<i>Judith A. Arter</i>	
Chapter 34	
Maybe We Learned All We Really Needed to Know in	
Kindergarten	
But How Could Anybody be Sure Until We Took the Test?	485
<i>Samuel E. Krug</i>	
Chapter 35	
Performance Assessment	
Designing Appropriate Performance Tasks and Scoring Rubrics	497
<i>Carole L. Perlman</i>	
Chapter 36	
Beyond Assessment to Best Grading Practice	
Practical Guidelines	507
<i>Laurie A. Carlson</i>	
Chapter 37	
Program Evaluation and Outcomes Assessment	
Documenting the Worth of Educational Programs	517
<i>Bradley T. Erford & Cheryl Moore-Thomas</i>	
Chapter 38	
Interpreting the Meaning of Test Results	
The Consultant's Role	525
<i>Donna M. Gibson</i>	

Chapter 39	
Testing FAQ	
How to Answer Questions Parents Frequently Ask About Testing ...	535
<i>Bradley T. Erford & Cheryl Moore-Thomas</i>	
Chapter 40	
Steps in the Right Direction	
Reporting Assessment Results to Students, Parents, School Board Members, and the Media	557
<i>Edward Roeber</i>	
Chapter 41	
Facilitating Career Development	
Assessment and Interpretation Practices	581
<i>Thomas F. Harrington & Richard W. Feller</i>	
Chapter 42	
Improving Work Life Decisions: O*NET™ Career Exploration Tools	595
<i>Phil Lewis & David Rivkin</i>	
Chapter 43	
Assessment of Workplace Stress	
Occupational Stress, Its Consequences, and Common Causes of Teacher Stress	611
<i>Jo-Ida Hansen & Brandon A. Sullivan</i>	
Section D. Musing Philosophical and Looking Toward the Future	
Chapter 44	
Current Issues in Educational Assessment	
The Test Publisher's Role	625
<i>William G. Harris</i>	
Chapter 45	
Technology, Collaboration, and Better Practice	
The Future of Assessment in Education and Counseling	649
<i>Jo-Ida Hansen</i>	

Chapter 46	
The Future of School Testing	
A School District Perspective	657
<i>Linda Elman</i>	
Chapter 47	
Harnessing the Power of Technology	
Testing and Assessment Applications	665
<i>Janet E. Wall</i>	
Chapter 48	
Computerized Adaptive Testing	
An Introduction	685
<i>Stephen G. Sereci</i>	
Chapter 49	
Assessing the Quality of Online Instruction	
Integrating Instructional Quality and Web Usability	
Assessments	695
<i>Anthony Ciavarelli</i>	
Chapter 50	
Needs Assessment	
An Ongoing Process for School Improvement	719
<i>Cheryl Moore-Thomas & Bradley T. Erford</i>	
Chapter 51	
The National Assessment of Educational Progress	
What It Tells Educators	729
<i>Lauress L. Wise</i>	
 Section E. Resources On Assessment	
Chapter 52	
The Joint Committee on Testing Practices	
Available Publications on Testing	745
<i>Lara Frumkin</i>	
Chapter 53	
Internet Resources in Educational Assessment	
A Webography	757
<i>Janet E. Wall</i>	

An Anthology of Assessment Resources769



Chapter 22

Test Consumers in the Military

Use of the Military Career Exploration Program in Schools

Janice H. Laurence

The military is not just a job—it's hundreds of jobs, with plenty of positions to boot. In terms of providing education, training, and employment, the military is unparalleled. The army, navy, marine corps, and air force enlist about 200,000 new recruits and commission more than 16,000 officers annually for active duty. These newcomers top off an incumbent strength of almost 1.4 million active members. Although most of the almost 900,000 selected reservists have had active duty experience, well more than 50,000 come in fresh from civilian life (Department of Defense, 2000).

Besides the traditional combat and seamanship roles, the enlisted military workforce comprises technicians, clerks, administrative associates, mechanics, computer specialists, high-tech equipment operators and repair specialists, health care specialists, and a host of other positions. Table 1 shows the occupational distribution of the enlisted ranks as of fiscal year 1999 (Department of Defense, 2000).

Table 1. Occupational Distribution of U.S. Military Enlisted Force (1999)

<u>Department of Defense Occupational Group</u>	<u>Percentage of Enlisted Force</u>
Infantry, gun crews, and seamanship specialists	17.0
Electronic equipment repair specialists	9.4
Communications and intelligence specialists	9.0
Medical and dental specialists	6.9
Other allied specialists	3.0
Functional support and administration	6.0
Electrical/mechanical equipment repair specialists	19.8
Craftspeople	3.5
Service and supply handlers	8.5
Nonoccupational military	6.9

About one in six enlisted members could be classified as a combat job incumbent or a general military employee, whereas one in five serves in a high-tech job in electronic equipment repair, communications and intelligence, or other allied specialist. Even combat jobs have become more technologically complex and relatively less labor intensive over the years—and more manpower has been added behind the combat scenes. Although most military jobs are in the blue collar category (infantry, gun crew, seamanship specialists; electrical and mechanical equipment repair specialists; and craftspeople), white collar positions (electronic equipment repair specialists; communications and intelligence specialists; medical and dental specialists; other technical and allied specialists; and administration) are almost as plentiful.

The most common jobs in the military are in electrical and mechanical equipment repair, with about one in five armed services workers engaged as an aircraft, automobile, and engine mechanic; ordnance mechanic; line installer; or radio, radar, and sonar equipment repair specialist. About one in six military workers is employed in administration as a stock and inventory clerk, shipping and receiving clerk, dispatcher, and the like.

The military services do not cull seasoned civilian workers to fill the ranks. Instead, they recruit novices and train them to perform myriad duties. Evidence shows that entry-level military jobs are more complex and demanding of workers than are civilian jobs (Laurence, 1994). Thus, selection and classification testing (i.e., assessment) is critical to staffing the military.

Military Career Counseling

Given military workforce requirements, is it any wonder that the military is a steadfast consumer—and producer—of career assessments? The military has in fact been a trailblazer with regard to cognitive test development and validation (Eitelberg, Laurence, & Waters, 1984). Numerous psychometricians and educational psychologists dedicate their efforts to maintain, update, advance, and monitor the exemplary cognitive testing program of the Department of Defense (DoD). The Armed Services Vocational Aptitude Battery (ASVAB; DoD, 1999) measures aptitudes in 10 areas (General Science, Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Numerical Operations, Coding Speed, Auto and Shop Information, Mathematics Knowledge, Mechanical Comprehension, and Electronics Information). Various combinations of these subtests are used to assess overall cognitive

aptitude as well as aptitudes for performing in specific jobs.

The ASVAB contributes to personnel selection and placement decisions and hence is an important component of military personnel readiness. The attention and resources focused on norming and validation with regard to technical training grades, administrative records, supervisory ratings, job knowledge test scores, and hands-on job performance measures are laudable and unparalleled (see, e.g., Bock & Mislevy, 1981; Fairbank et al., 1990; Green & Wigdor, 1991; Green, Wing, & Wigdor, 1988). Indeed, ASVAB results reliably indicate one's standing relative to the U.S. population of youth ages 18 to 23. Time and again, studies have shown that subtest composite scores fairly and validly assess the likelihood of achieving technical proficiency or effectiveness across the wide spectrum of jobs found in the military (most of which have civilian counterparts).

Since 1968, the DoD has offered the ASVAB at no cost to high schools nationwide to promote career exploration and to facilitate recruiting. Known originally as the Student Testing Program (STP), this idea blossomed over the years into the Career Exploration Program (CEP)—a professional and comprehensive career counseling tool for schools and students. Service recruiters receive the names and ASVAB scores of participating students who agree to have this information released. Thus, there are strings attached to CEP participation, but they are not demanding.

Each year, about 900,000 students in more than 14,000 schools take the ASVAB. More than one fourth of high school seniors participate in the CEP at some point during high school (Baker, 2000). The CEP is designed to help students, primarily 11th- and 12th-graders, explore both military and civilian careers through materials that support educational and career counseling. Recruiters can use the results to identify individuals who qualify for military service. Three primary CEP components assess aptitudes, interests, and work values:

1. The 10 ASVAB subtests are combined and scores are reported on three composites: Verbal Ability, Math Ability, and Academic Ability. ASVAB codes highlight similarities between the aptitude levels of test takers and those of incumbents already performing various jobs. Military Career Scores estimate the likelihood that an individual will qualify for enlistment.
2. The Interest-Finder identifies areas of interest to the test taker (Realistic, Investigative, Artistic, Social, Enterprising, and Conventional).

3. OCCU-FIND links ASVAB and Interest-Finder results, along with other information (e.g., educational goals, work values) to 201 occupations organized by interest area.

Detailed test results (and interpretation) are provided to students, with copies for counselors. Besides the support provided by Education Services Specialists (ESS), civilians with an educational or counseling background, and recruiters, materials are available to help school staff, students, and their parents get the most out of the CEP. These include the *Educator and Counselor Guide*, *Student and Parent Guide*, *Counselor Manual*, *Student Workbook*, *Military Careers*, *Technical Manual*, and *Recruiter Guide*. (Most of these documents are available for download from the ASVAB website at www.asvabprogram.com.) The ASVAB is also incorporated into many Career Information Delivery Systems (CIDS)—computerized career information systems made available by states, regions, and commercial vendors.

Recruiting

Military recruiting is always challenging. Getting the word out about military career opportunities is therefore a vital service of the CEP. The ASVAB CEP is an effective marketing and recruiting tool. The program is valued by recruiters as a means of obtaining access to schools, making contact with individual students, and identifying those who are qualified for and interested in military service. Up to one fifth of CEP participants subsequently enlist in the military (Laurence & Ramsberger, 1999).

Evidence suggests that the CEP is a positive influence on those who formerly held neutral or negative views regarding military service (Laurence, Wall, Barnes, & Dela Rosa, 1998). CEP participants are more likely to express an interest in joining one of the military services as a result of the information obtained through the CEP. In addition, data suggest that CEP participants are more likely than nonparticipants to view the military as a place where they can obtain money for education, learn a valuable trade or skill, and receive job preparation.

The ASVAB CEP targets non-college-bound youth. Largely because of its vocational emphasis, the CEP has traditionally been more attractive to young people who are not considering postsecondary education, at least not for the immediate future. Given the increasing numbers of students choosing postsecondary educational opportunities, however, it is important for students to recognize the college

opportunities afforded by the military, such as the Voluntary Education Program, the Montgomery GI Bill, and the officer track (Asch, Kilburn, & Klerman, 1999). Besides exploring career and other opportunities afforded by the military, college-bound youth can benefit from exposure to the CEP testing process and outcomes.

Career Decisions

Schools that participate in the CEP choose to do so for a number of reasons: the program is free; it is an effective tool for counseling non-college-bound youth; it provides an opportunity for military career exploration; and it is a readily available, well-documented career exploration tool. Further, the CEP is comprehensive and effective in meeting school career counseling needs, has a positive impact on student career exploration, and is at least as good as other programs (Laurence & Ramsberger, 1999). The vocational emphasis of the program as well as the supplementary materials (e.g., *Student Workbook*) and counseling support provided by the military fill a void, especially in economically deprived schools. Although many students are well prepared for the frenetic activities of registering, paying, and convening for the ACT Assessment or the SAT, others, without plans for college or mentors to show them the ropes, might well remain forgotten without the CEP.

The ASVAB alone provides invaluable information for civilian career counseling. Composites from the ASVAB are predictive of high school course grades (Fairbank, Welsh, & Sawin, 1990). ASVAB tests also correlate highly with comparable tests from civilian aptitude and achievement batteries (Department of Defense, 1999). Based on patterns of ASVAB scores, Armstrong, Chalupsky, McLaughlin, and Dalldorf (1988) classified a sample of individuals into their civilian occupations with a statistically significant degree of accuracy. Even more salient is a study that provides direct evidence of the criterion-related validity of the ASVAB for a sample of 11 different civilian occupations (e.g., bus driver, computer operator, word processor, nurse, electronics technician; Holmgren & Dalldorf, 1993). Further, the accepted theory of validity generalization together with the results of a military-civilian occupational crosswalk extend this mound of evidence from military occupations and the congruent findings from selected civilian jobs to additional occupations. In other words, the ASVAB has demonstrated validity for military and civilian jobs. It is technically acceptable to extrapolate these findings to encompass jobs for which performance is validly predicted by measures highly correlated with ASVAB and for

jobs that are highly similar to those included in ASVAB validation studies (Department of Defense, 1999). That is, there is sound statistical evidence that test validity is not situation- or job-specific; rather, if validity is established in one job, it holds for similar jobs. Certainly ASVAB validity has been established above and beyond applicable professional testing guidelines and practices.

The DoD has gone beyond investment in the development and administration of the ASVAB, and program evaluation extends beyond its value in recruiting. Systematic evaluation efforts have provided sound evidence that adolescents who participate in this broad-based program show an increase in career development efforts (Baker, in press; Levine, Huberman, & Wall, 1996). The national normative base of 18- to 23-year-olds, most appropriate for enlistment decisions, was supplemented for CEP use with a high school sample of almost 10,000 students in grades 10 through 12. The inclusion of the additional sample of high school students reinforces the utility of the CEP, especially for participants in 10th grade (Department of Defense, 1999).

The CEP is based upon sound psychometric and vocational personality theory (Wall & Baker, 1997). Participants are provided with more than just scores indicating their standing relative to others; the program helps students to identify occupations consistent with their interests, abilities, and values. The program provides practical information regarding the cognitive demands of and typical educational preparation needed for particular jobs, and the degree to which these jobs match one's preferences for certain activities and the values that one is looking to satisfy through one's career (e.g., challenge, creativity, physical activity, independence; Wall, 1994). This comprehensive and integrated program under DoD's aegis promotes knowledge of self, occupational opportunities, and the world of work. It reduces career confusion and facilitates judgments of career attractiveness (Baker, in press).

Some Parting Thoughts on the CEP

With its dual goals of recruiting and career counseling, the CEP does not operate without suspicion or conflict. Those suspicious of military recruiting efforts can rest assured that the program has strong technical underpinnings. Aptitudes, interests, and preferences are indeed linked to civilian, not just military, jobs. Occupations included for exploration in the OCCU-FIND represent "the range of diversity in the world of work" (Wall, 1994, p. 610). Rather than limiting options, the

CEP encourages rather wide and warranted exploration. The accompanying materials highlight occupations within two contiguous cognitive complexity levels, three interest areas, and up to six personal preferences (Wall, Wise, & Baker, 1996). Certainly, an aim of the program is to garner recruiting leads; however, participants may opt not to share their results with military recruiters.

There is conflict with regard to participation because military recruiters would prefer to test only high school seniors—those who have a shot at helping them meet their recruiting objectives. There is no outcry at including juniors, but extending the CEP to sophomores (or freshmen) may be viewed as a waste of precious recruiting resources and detrimental to recruiters' short-term, "put 'em in boots" perspective. Needless to say, from a career counseling perspective, career exploration should begin early—well before the senior year of high school. This conflict does not speak ill of the program. Quite the contrary; it is the effectiveness of the CEP for recruiting and career counseling that is at the conflict's core.

Although the program is already top-notch, improvements are on the horizon. In response to demographic trends and changes in the workplace, DoD is modifying its testing and assessment practices and technical underpinnings. The psychometric properties and functioning of the Interest-Finder are scheduled for a tune-up as are the ASVAB's accompanying materials. What's more, the version of the ASVAB that is used for operational enlistment decisions is expected to have an interest measure folded in before long.

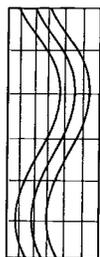
The military offers education, training, and employment to novices to the workforce, our nation's youth. The military continues to be a trailblazer with regard to testing and human resource assessment. No compendium on career counseling would be complete without mentioning the military. This chapter provides merely a condensed snapshot of the CEP and DoD's commitment to career assessment for both military and civilian careers.

References

- Armstrong, T. R., Chalupsky, A. B., McLaughlin, D. H., & Dalldorf, M. R. (1988). *Armed Services Vocational Aptitude Battery: Validity for civilian occupations* (AFHRL-TR-88-20). Brooks Air Force Base, TX: Air Force Human Resources Laboratory.

- Asch, B. J., Kilburn, M. R., & Klerman, J. A. (1999). *Attracting college-bound youth into the military*. Santa Monica, CA: RAND.
- Baker, H. E. (2000, July). *Cross-sectional vs. cohort participation rates*. Paper presented at the meeting of the National Center for Education Statistics (NCES) Summer Data Conference, Washington, DC.
- Baker, H. E. (in press). Reducing adolescent career indecision: The ASVAB Career Exploration Program. *Career Development Quarterly*.
- Bock, R. D., & Mislevy, R. J. (1981). *The profile of American youth: Data quality analysis of the Armed Services Vocational Aptitude Battery*. Chicago, IL: National Opinion Research Center.
- Department of Defense. (1999, December). *Technical manual for the ASVAB 18/19 Career Exploration Program* (DoD 1304.12-LASTP-TS). North Chicago, IL: HQ USMEPCOM.
- Department of Defense. (2000, November). *Population representation in the military services: Fiscal year 1999*. Washington, DC: Office of the Assistant Secretary of Defense (Force Management Policy).
- Eitelberg, M. J., Laurence, J. H., & Waters, B. K. (1984). *Screening for service: Aptitude and education criteria for military entry*. Washington, DC: Office of the Assistant Secretary of Defense (Manpower, Installations, and Logistics).
- Fairbank, B. A., Welsh, J. R., & Sawin, L. L. (1990). *Armed Services Vocational Aptitude Battery (ASVAB): Validity of ASVAB 14 for the prediction of high school course grades* (AFHRL-TR-90-48). Brooks Air Force Base, TX: Air Force Human Resources Laboratory.
- Green, B. F., Jr., & Wigdor, A. K. (1991). Measuring job competency. In A. K. Wigdor and B. F. Green Jr. (Eds.), *Performance assessment for the workplace* (Vol. II, pp. 53-74). Washington, DC: National Academy Press.
- Green, B. F., Jr., Wing, H., & Wigdor, A. K. (Eds.). (1988). *Linking military enlistment standards to job performance: Report of a workshop*. Washington, DC: National Academy Press.

- Holmgren, R. L., & Dalldorf, M. R. (1993, October). *A validation of the ASVAB against supervisors' ratings in civilian occupations*. Palo Alto, CA: American Institutes for Research.
- Laurence, J. H. (1994). *The military: Purveyor of fine skills and comportment for a few good men* (EQW Catalog Number WP25). Philadelphia, PA: University of Pennsylvania's Center for the Educational Quality of the Workforce.
- Laurence, J. H., & Ramsberger, P. F. (1999, October). *Evaluation of the DoD Armed Services Vocational Aptitude Battery Career Exploration Program* (FR-WATSD-99-46). Alexandria, VA: Human Resources Research Organization.
- Laurence, J. H., Wall, J. E., Barnes, J. D., & Dela Rosa, M. (1998). Recruiting effectiveness of the ASVAB Career Exploration Program. *Journal of Military Psychology, 10*(4), 225-238.
- Levine, R., Huberman, M., & Wall, J. (1996, July). *ASVAB Career Exploration Program: Impact on student career development* (Technical Report 96-006). Seaside, CA: Defense Manpower Data Center.
- Rohrback, M., Barnes, J. D., Laurence, J. H., & Wall, J. (1996, September). *ASVAB Career Exploration Program: Impact on military recruiting* (Technical Report 96-008). Monterey, CA: Defense Manpower Data Center.
- Wall, J. E. (1994). An example of assessment's role in career exploration. *Journal of Counseling & Development, 72*(6), 608-613.
- Wall, J. E., & Baker, H. E. (1997). The Interest-Finder: Evidence of validity. *Journal of Career Assessment, 5*(2), 255-273.
- Wall, J. E., Wise, L., & Baker, H. E. (1996). Development of the Interest-Finder: A new RIASEC-based interest inventory. *Measurement and Evaluation in Counseling and Development, 29*, 134-152.



Chapter 23

Educational Assessment in a Reform Context

Michael H. Kean

It is difficult to believe that there was a time when the day's news *didn't* contain any mention of educational assessment. Up until the late 1980s, governors, state legislators, members of Congress, journalists, and other pundits knew little and said even less about how U.S. students were measured and educational programs were evaluated. Although educational assessment has played a pivotal role in American education for well more than 50 years, it remained in the background of our nation's policy debates and was considered a technical, if not esoteric, field.

Fast forward to today. Governors' speeches are peppered with remarks about accountability and standardized testing. Members of Congress engage in lengthy and often acrimonious debate over proposals for national testing of elementary and secondary school students. Local journalists routinely report on educational standards and testing. Moreover, the discussions do not end in the political arena. In political polling parlance, testing has become nightly "table talk" over dinner for moms, dads, and their kids.

At the same time (and through no coincidence), assessment is playing greater roles in the current educational environment. Assessment results are a major force in shaping public perceptions about the achievement of our students and the quality of our schools. Educators use assessment results to help improve teaching and learning as well as to evaluate programs and the effectiveness of schools. Educational assessment is also used to generate the data on which policy decisions are made. Because of the important role it performs, assessment is a foundational activity in every school, in every school district, and in every state.

What events and trends led to the transformation of educational assessment into nightly table talk? Why is there now a strong political dimension to educational assessment? Which aspects of assessment should educators and policymakers bear in mind as they go about their

work? This chapter will provide answers to all these questions.

The Political Context of Reform

Over the past 20 years, education reforms have generally been of three types: structural, process, or content:

Structural reform refers to changes in the *structure* of education, such as a longer school day or school year, smaller class sizes, magnet schools, charter schools, or a middle school versus junior high school system.

Process reform refers to the *way* in which teachers teach and students learn. Team teaching, reading recovery, and use of educational software are examples of process reform.

Content reform refers to *what* teachers teach. Examples are phonics or whole language approaches to reading, new math, and standards-based curricula.

Testing entered the political realm with the advent of standards-based school reform, which is both a process reform and a content reform. This reform focuses on improving our schools, increasing student achievement, and building accountability for results through a system with three primary components: (a) new (and higher) *standards*, (b) new *assessments* designed to measure those standards, and (c) *consequences* for meeting or not meeting the standards. Politics is part of this process because of its traditional and rightful (but often unpredictable) role as the driver of policy in our national and state democracies.

The standards-based movement emerged in the early 1990s as a response to the call to arms issued by the 1983 release of *A Nation at Risk*. This slim but seminal report from the National Commission on Excellence in Education characterized U.S. schools as wholly inadequate and went so far as to say, "If an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war" (p. 1). In short order *A Nation at Risk* galvanized policymakers at the federal and state levels. The nation's governors, acting collectively through the National Governors Association, developed and issued *Time for Results* (1986), a report that called for, among other things, greater

accountability in our nation's public schools. Out of this period emerged a group of "education governors" who would later make their mark in education on the national scene: Tennessee Governor Lamar Alexander, Arkansas Governor Bill Clinton, South Carolina Governor Richard Riley, and Colorado Governor Roy Romer. Whereas Alexander and Riley would serve as U.S. Secretaries of Education in the 1990s and Romer would lead many national panels on education, Clinton forged a legacy as the nation's most active education president.

By 1989, concern over the nation's schools reached the level where the governors and President George H. Bush convened the first ever National Education Summit, in order to propose solutions. The fall summit, held at the University of Virginia in Charlottesville, culminated in an agreement to set six (later expanded to eight) broad National Education Goals. The goals were developed and released in 1990. At the same time a federal commission—The National Education Goals Panel—was created by Congress to track national and state efforts to reach these goals by the year 2000.

Although the National Education Goals were not reached by 2000, their impact was felt in two ways. First, they focused public attention on the need for increased student achievement. Second, they served as the starting point for the development of new education standards. This development began at both the federal and state levels, though it was action at the federal level that spurred many states to begin developing and setting their own standards.

Federal action came initially in April 1991 in the form of America 2000, the George H. Bush administration's education proposal. America 2000 set forth voluntary national standards in a range of subject areas and proposed a series of national tests. Although America 2000 did not find its way into law by the end of the first Bush administration, the Clinton administration came forward with a similar proposal, called Goals 2000 (signed into law as the Goals 2000: Educate America Act in 1994). America 2000 and Goals 2000 had some distinct differences, but they were alike in their drive for high standards and new assessments to measure student progress.

Goals 2000 became the most pervasive national K–12 education policy in a generation. It provided federal incentives for states to create new systems of accountability by setting their own standards and creating new assessments, which the states did. At the start of the decade only a handful of states had academic standards. By the end of it, close to 50 states had developed standards.

Despite its pervasiveness and its affinity to the America 2000

proposal, Goals 2000 found itself in the mid- to late-1990s under increasing attack from Republicans and conservatives, who felt the federal government had overextended its reach into state and local education policies. Republican critics claimed that while Washington had historically funded K–12 education at low levels (current funding is approximately nine cents on the dollar), it exerted too much authority in local classrooms. This sentiment led to a policy standoff in the fall of 1997 when the Clinton administration watched its proposal for voluntary national tests in reading and mathematics go down to defeat on Capitol Hill.

In early 2001, the administration of President George W. Bush introduced No Child Left Behind as its proposal for the reauthorization of the Elementary and Secondary Education Act (ESEA). The legislation sought greater accountability through annual testing in grades 3 through 8 in reading and mathematics, but left states to set their own standards and choose their own tests. In doing so, Washington not only re-established individual student progress as a central tenet of ESEA, it also found a politically acceptable compromise on assessment. In late December 2001, Congress passed the legislation by a wide bipartisan margin. President Bush signed the act into law soon after.

Federal Policy Issues

Invariably, and sometimes unfortunately, a recurring set of issues continues to evolve around Washington education debates. Like entrenched armies on the Western Front in the First World War, politicians often fight battles over and over on the same ground for years, and no real victor emerges. Typically, education debates in Washington have to do with the federal government's regulatory power and its authority over our nation's decentralized public education system.

Washington's authority. Local and state control of education is a deeply rooted concept in the United States. It remains so today, with the 50 states and tens of thousands of localities providing 91 percent of the funding at the K–12 level. Not a single education bill is debated in Congress today without at least one lawmaker (and usually many more) questioning the authority of Washington to impose educational mandates on the states and the nation's 15,600 school districts. Lawmakers from both sides of the aisle raise the issue, particularly when Congress mandates billion-dollar programs such as the Individuals With Disabilities Act (IDEA) and fails to fully fund them.

The “devolution revolution.” With the Republican sweep of Capitol Hill in the November 1994 midterm elections, the devolution revolution was set in motion. The idea is to devolve as much federal authority as possible to the states and localities, where better decisions might be made. Although this devolution is often viewed as a Republican philosophy, many centrist Democrats also favor devolution initiatives. To date, the revolution has been seen most clearly in Congress’s massive overhaul of welfare and job training programs. It has also appeared in the education arena, however, where it emerges in debates over block grants, program consolidation, and “ed flex,” all of which opt to lift regulations prohibiting the blending of federal dollars from various programs. Whereas Democrats argue that federal education programs and their accompanying dollars should be carefully targeted to specific populations, Republicans counter that regulations should be lifted so that states and local schools can determine how best to use federal monies. There has not been a clear winner in the debate. Although more flexibility has been provided in various laws, many federal programs—rightly or wrongly—remain prescriptive in their aims and targeted populations.

Testing on a national scale. Between 1991 and 2001 Congress has had three major debates over testing: first with America 2000 in 1991, second in 1997 with the Clinton administration’s voluntary national test proposal, and again in 2001 with the testing proposal in No Child Left Behind. Each debate has raised concerns over Washington’s role in dictating how states should evaluate students.

Opportunity to learn standards. In 1991 congressional critics of America 2000 argued that if Washington was going to require new, higher academic standards, schools should have increased funding so that they could better prepare students to reach those standards. This same argument has emerged in 2001 as Congress debated the reauthorization of the ESEA.

Use of the National Assessment of Educational Progress (NAEP). Congress has often debated the notion of expanding the NAEP to measure individual student progress. Historically, the NAEP mission has been to intermittently sample student performance in various subject areas. Because of that, various attempts to expand NAEP have encountered opposition on Capitol Hill from lawmakers who fear NAEP expansion would lead to a national test.

State Policy Issues

The list of issues at the state level is more extensive than at the federal level because state policymakers, unlike members of Congress, have been closely involved in setting standards and shaping new assessment programs. The list of important issues for these policymakers ranges from the use of multiple measures to legal defensibility to public relations.

Governance. Consideration of major public policy in any state is a complex undertaking involving a number of different policymakers. While many governors play a central role in leading education reform in their states, at least three other individuals or entities—the state commissioner or superintendent of education, the state board of education, and the state legislature—play crucial roles, too. As the “dance of legislation” occurs, each of these individuals and entities contribute to the debate in some way.

The need for the right kind of information. States are in a unique position to use assessments for generating the types of data that policymakers, educators, and parents need to make decisions about their schools and students. State assessments more frequently serve as the “accountability fulcrum.” Why? Because most assessment programs at the local school district level are designed primarily to improve teaching and learning, not to collect extensive, reliable data on student performance. Meanwhile, at the national level, the NAEP—an assessment sanctioned and funded by the federal government—generates snapshots of how small samples of students are performing in a given subject at a particular grade level. NAEP cannot expand on this snapshot function without igniting debates on federal versus state and local governance in education. This situation provides the states with the opportunity to generate more relevant statewide data on their students and school systems. Typically, this is accomplished by giving students a standardized, norm-referenced test. This type of test yields a variety of rich, reliable data that can be used for both statewide accountability purposes and to determine individual pupil progress toward meeting state standards.

Sequencing. Successful standards-based reform is based on a sequence where goals are developed first, followed by standards, then new curricula and instructional approaches, and finally assessments.

Standards setting. How standards are developed has been a very important issue for states. Great care must be taken to ensure that educators, policymakers, business leaders, and other key players are involved in creating new standards. Many states have developed both curriculum standards—criteria describing what students should know and be able to achieve—and performance standards—levels of acceptable student performance. Major test publishers work with standards-setting groups in the states to ensure that newly formulated standards can be measured by valid, reliable, and fair assessments.

High-stakes testing. The term *high stakes* refers to the use of assessments for purposes such as promotion and retention of students; graduation or exit exams; and rewards or penalties for schools or educators based upon student performance.

Legal defensibility. Because of the trend toward high-stakes testing, legislators and other state policymakers must ensure that state testing programs can withstand legal challenges. For instance, the number of lawsuits over high school graduation exams is increasing and is likely to continue to do so. Because of that, states must work closely with their assessment contractors to see that the tests used are valid, fair, and reliable.

Multiple measures. No single test can do it all, and no single assessment should serve as the sole evaluation tool in measuring performance. Multiple measures such as additional tests, grades, and teacher-made classroom quizzes—must be used to fully gauge student achievement.

Inclusion. Standards-based education reforms aim to set higher expectations for all students. In doing so, however, great care must be taken to accommodate children with special needs and those whose first language is not English. These children must not be left behind. The very core of standards-based reform is opportunity: the opportunity for *all* children to learn. Legislators, state education departments, curriculum developers, and test and textbook publishers have moved quickly in recent years to ensure that all students have the tools they need to learn and to demonstrate their knowledge and skills.

Report cards. Nearly all states (and many local school districts) now publish and disseminate report cards on individual districts and schools. These report cards serve a valuable function in informing parents and

the public about the performance of their local public schools.

Communications and public relations. The standards-based movement represents a very significant change in how our schools go about educating children. Students, entire schools, and in some cases, teachers and administrators, must now meet higher expectations. When they do not, there may be consequences. Students may not be allowed to graduate, schools may face reconstitution with new leadership and teachers, and teachers may face loss of merit pay. Listing these consequences is not meant to cast standards-based reform in a negative light. Reform has generally been successful in bringing to public schools new standards of excellence, innovative curricula, challenging assessments, and new teaching strategies. Unfortunately, not all members of the public see and understand these positive changes. They see only the bad news of the high-stakes era. We now know why. Well into the 1990s, educators, policymakers, and the schools failed to educate the public about standards-based reform. Whereas some key audiences, such as the business community, were brought on board early, many parents still do not understand the need to hold students to higher expectations through new standards and assessments. As a result, a small but shrill cadre of testing critics has created a testing backlash in some communities. Although this backlash is unlikely to do serious harm to the standards-based reform movement, it represents a lesson policymakers should heed: *Always communicate (and keep communicating) the benefit of your reforms to key audiences.* Use public relations strategies to build understanding and support for reform among teachers, parents, students, and the community at large.

A Final Word: The Second Decade

The various federal and state issues outlined here represent the *current* political context that surrounds educational assessment. We are now in the second decade of standards-based education reform. Like any significant public policy change, the reform movement will be modified and refined in coming years. Educators and public policymakers should anticipate the debates that lead to these refinements. Crystal balls are usually murky at best, but we can anticipate the following changes:

The Bush administration's annual testing initiative. The first challenge will be to coalesce local and state testing programs in ways

that meet the Bush administration proposal for annual testing of students in mathematics and reading in grades 3 through 8. Great care must be taken to respect state and local educational goals and curricula while establishing annual testing regulations. Any suggestion from Washington regarding the shape or content of assessments could very well lead to resistance or a backlash from governors and state policymakers.

Revised standards. Educational standards are not static. They must evolve based on society's needs to educate and train its children. This means educators and policymakers must continue to research, write, and rewrite state and local standards. In this process public debates will occur over the content of standards and over how high to set standards at particular grades. It will also mean that classroom curricula, teacher training programs, and assessments will undergo constant modifications to reflect these new standards.

The blending of curriculum and assessment. Curriculum developers, educational technologists, and textbook and test publishers are working diligently to bring innovations to the classroom. Over the course of the next 10 to 20 years the greatest advance in standards-based education may be the blending of curriculum and assessment. Test items will be embedded in educational software so students can be measured as they learn. As a result, evaluation will become transparent and less time will be spent on taking formal tests. This development, perhaps more than any other, will silence the critics of assessment and cause the testing backlash to melt away.

An educated public. Despite the failure of the standards-based movement to quickly educate parents about standards, assessments, and high-stakes consequences, new public information campaigns will be designed to reach out to all sectors of the public and to build greater understanding and support.

Teaching oriented to standards. In the same way that the public was left behind in the first decade of the reform, so were many teachers who were not trained to teach to specific state standards. However, new teacher training programs for college students and in-service programs for current teachers are beginning to create a new cadre of educators oriented to standards-based reform.

Finally, if the reform movement is to reach its true potential, the second decade must be one in which no segment of the public is forgotten. Everyone engaged in public education—governors, parents, children, teachers, boards of education, school superintendents, state legislators, publishers, researchers, school administrators, college faculty, teachers' unions, and members of Congress—should have permanent seats at the table of education reform. The creation of public policy requires the firm and active participation of all affected publics. Education reform and assessment will always have its political and policy dimensions, but the inclusion of all publics will provide a firm foundation upon which to build such reforms.

References

- National Commission on Excellence in Education (1983, April). *A Nation at Risk*. Washington, DC: U.S. Department of Education.
- National Governors Association. (1986). *Time for Results*. Washington, DC: Author.
- U.S. Department of Education. (1991). *AMERICA 2000: An Education Strategy*. Washington, DC: Author.
- ◆ U.S. Department of Education. (2001). *No Child Left Behind*. Washington, DC: Author.

◆ Document is included in the Anthology of Assessment Resources CD