IAP  AP 101

APPLIED PSYCHOMETRICS 101:

#9: The Armed Services Vocational Aptitude Battery (ASVAB): Why it should not be used to in the determination of a diagnosis of mental retardation / intellectual disability

ASVAB scores are often incorrectly interpreted as a measure of general intellectual functioning in the context of determining if a person is an individual with (or without) mental retardation (MR) / intellectual disability (ID). The ASVAB is an aptitude battery and not an intelligence test battery. Although the differentiation between aptitude and intelligence test batteries measures may (at times) sound fuzzy, the distinction between the two is critically important, particularly regarding how the different respective batteries are designed, the abilities they each measure, and how the resultant scores should be validly interpreted. The aptitude -intelligence test battery distinction is clearly defined in psychological measurement fields. Although aptitude and intelligence batteries often measure some overlapping abilities, the ASVAB-as-an-aptitude measure is often confused with the incorrect interpretation of the ASVAB-as-IQ (general intelligence) measure. This report explains the distinction.

Kevin S. McGrew, Ph.D.

Educational Psychologist

Director

Institute for Applied Psychometrics (IAP), LLC

© Institute for Applied Psychometrics $_{llc}$ (IAP) 11-18-2010 ver1

Definition and purpose of the ASVAB and aptitude tests

According to the American Psychological Association's Dictionary of Psychology, the ASVAB is a *group-administered* paper and pencil test battery:

> "developed in 1966 by the Department for Defense for use by the U.S. military as a standardized instrument for *personnel selection and classification (specific job assignment*)"

APA Dictionary of Psychology, American Psychological Association 70 (Gary R. Vandenbos ed., 2007) (emphasis added).

ASVAB tests are selected based on their *degree of perceived similarity to military occupations*, in order to produce a battery with *good prediction, selection and classification of performance or trainability in the military.* See A. Anastasi & S. Urbina, Psychological Testing (7th ed. 1997); R. Gregory, Aptitude Tests, in The Encyclopedia of Intelligence 110-117 (R. Sternberg, ed., 1994); A.R. Jensen, The g Factor: The Science of Mental Ability (1998); R.D. Roberts, G.N. Goff, F. Anjoul, P.C. Kyllonen, G. Pallier & L. Stankov, The Armed Services Vocational Batter (ASVAB) – Little More Than Acculturated Learning (Gc)!?, 12(1) Learning and Individual Differences 81-103 (2000); J. Welsh, T. Watson & M. Ree, Armed Services Vocational Aptitude Battery (ASVAB): Predicting Military Criteria From General and Specific Abilities (AFHRL-TR-90-22) Brooks, AFB, TX: U.S. Air Force Human Resources Laboratory (1990).

Definition and purpose of tests of general intelligence (IQ)

According to the American Psychological Association's Dictionary of Psychology, an intelligence test is:

> "…an *individually administered*, standardized test used to determine a person's level of intelligence by measuring his or her *ability to solve problems, form concepts, reason, acquire detail, and perform other intellectual tasks."*

APA Dictionary of Psychology, American Psychological Association 70 (Gary R. Vandenbos ed., 2007) (emphasis added).

According to the American Association on Intellectual and Developmental Disabilities (AAIDD) official manual for defining and classifying intellectual disability, *intelligence* is defined as follows:

> "For purposes of diagnosis, intellectual functioning is currently best conceptualized and captured by a general factor of intelligence. Intelligence is a general mental ability. It includes reasoning, planning, solving problems, thinking abstractly, comprehending simple ideas, learning quickly, and learning from experience. The "significant limitations in intellectual functioning" criterion for a diagnosis of intellectual disability is an IQ score that is approximately two standard deviations below the mean, considering the standard error of measurement for the specific instruments used and the instruments strengths and limitations."

See American Association on Intellectual and Developmental Disabilities, Intellectual Disability: Definition, Classification, and Systems of Supports 31 (2010).

Scientific and professionally recognized contemporary measures of general intelligence (e.g., WAIS-III/WAIS-IV; Stanford-Binet IV; Woodcock-Johnson Battery III) have as their *primary design objective* to include *a representative sampling of an array of mental abilities that constitute the recognized domain of general intelligence (g).* See R. Gregory, Aptitude Tests, in The Encyclopedia of Intelligence 110-117 (R. Sternberg, ed., 1994); R. Snow, Abilities and Aptitudes, in Encyclopedia of Intelligence 3-5 (1994).

Primary difference between ASVAB (aptitude) and intelligence (IQ) tests

Although the differentiation between aptitude and *intelligence* test batteries measures may (at times) sound fuzzy, the distinction between the two is critically important, particularly regarding how the different respective batteries are designed, the abilities they each measure, and how the resultant scores should be validly interpreted. See R. Gregory, Aptitude Tests, in The Encyclopedia of Intelligence 110-117 (R. Sternberg, ed., 1994). The aptitude/intelligence test battery distinction is clearly defined in psychological measurement fields. Although aptitude and intelligence batteries often measure some overlapping abilities, the *ASVAB-as-an-aptitude measure* is often confused with the incorrect interpretation of the *ASVAB-as-IQ (general intelligence) measure.*

The primary purpose and design of contemporary individually-administered *intelligence (IQ)* test batteries is to provide the best possible estimate of "*general intelligence (g)" by sampling from the major domains of human intelligence* (viz., Cattell-Horn-Carroll theory of cognitive abilities; CHC theory). See A. Kaufman, IQ Testing 101 (2009); K. McGrew, Analysis of the Major Intelligence Batteries According to a Proposed Comprehensive Gf–Gc Framework, in Contemporary Intellectual Assessment: Theories, Tests, and Issues 151-179 (1997); K. McGrew, The Cattell–Horn–Carroll Theory of Cognitive Abilities, in Contemporary Intellectual Assessment: Theories, Tests, and Issues 136-181 (2005); K. McGrew, Editorial, CHC Theory and the Human Cognitive Abilities Project: Standing on the Shoulders of the Giants of Psychometric Intelligence Research, 37 Intelligence 1-10 (2009).[1]

The research evidence indicates that of the major individually administered comprehensive IQ test batteries with adult norms (WAIS-III/WAIS-IV, Stanford-Binet V, Woodcock Johnson Battery III), these batteries include tests draw from 4 to 7 of the broad CHC broad ability domains.[2] The domains include

[1] CHC theory is the contemporary name for this theory/model. It also continues to frequently be referred to, in the professional and research literature, as the Horn-Cattell Gf-Gc theory, Gf-Gc theory, or the Expanded/Extended Gf-Gc theory.

[2] The classification of CHC abilities measured by contemporary IQ tests referenced in this statement is based on in the following CHC-based intelligence testing research literature: D.P. Flanagan, Wechsler-Based CHC Cross-Battery Assessment and Reading Achievement, 15(3) School Psychology Quarterly 295-329 (2000); D.P. Flanagan, S. Ortiz, & V.C. Alfonso, Essentials of Cross-Battery Assessment (2nd Ed. 2007); D. Flanagan & K. McGrew, Interpreting Intelligence Tests From Modern *Gf-Gc* Theory - Joint Confirmatory Factor Analysis of the WJ-R and Kaufman Adolescent and Adult Intelligence Test in a Non-White Sample, 36 Journal of School Psychology 151-182 (1997); K.S. McGrew, Analysis of the Major Intelligence Batteries According to a Proposed Comprehensive Gf–Gc Framework, in Contemporary Intellectual Assessment: Theories, Tests, and Issues 151−179 (D. P. Flanagan, J. L. Genshaft, & P. L. Harrison, eds., 1997); K.S. McGrew, The Cattell–Horn–Carroll Theory of Cognitive Abilities, in Contemporary Intellectual Assessment: Theories, Tests, and Issues 136−181 (D. P. Flanagan & P. L. Harrison, eds., 2005); K.S. McGrew & D.P. Flanagan, The Intelligence Test Desk Reference (ITDR): Gf–

verbal abilities or crystallized intelligence (Gc), fluid reasoning or intelligence (Gf), short-term and working memory (Gsm), visual-spatial processing (Gv), processing speed (Gs), quantitative knowledge/reasoning (Gq), long-term storage and retrieval (Glr), and auditory processing (Ga). The representativeness and breadth of sampling from the primary abilities found in the domain of general intelligence is one of the primary criterion (that of "content and construct validity," see American Educational Research Association, American Psychological Association, and National Council on Measurement, Standards for Educational and Psychological Testing (1999) (also known as the Joint Test Standards)) for the interpretation of the global IQ scores as representing general intelligence (g).

In contrast, the test design and selection procedures used to develop the initial and subsequent editions of the ASVAB[3] were (and continue) to focus on designing a battery with good *prediction, selection and classification of performance or trainability in the military*, in order to assign recruits to appropriate jobs and training programs. See A. Anastasi & S. Urbina, Psychological Testing (7th ed. 1997); R. Gregory, Aptitude Tests, in The Encyclopedia of Intelligence 110-117 (R. Sternberg, ed., 1994); R.D. Roberts, G.N. Goff, F. Anjoul, P.C. Kyllonen, G. Pallier & L. Stankov, The Armed Services Vocational Batter (ASVAB) – Little More Than Acculturated Learning (Gc)!?, 12(1) Learning and Individual Differences 81-103 (2000); J. Welsh, T. Watson & M. Ree, Armed Services Vocational Aptitude Battery (ASVAB): Predicting Military Criteria From General and Specific Abilities (AFHRL-TR-90-22) Brooks, AFB, TX: U.S. Air Force Human Resources Laboratory (1990).

Predicting (forecasting) a person's probability of success in learning knowledge and skills in specific and narrow educational, occupational, or military settings is the classic design objective of an *aptitude* battery such as the ASVAB, college admission tests (e.g., SAT, ACT), admission tests for postgraduate professional training (e.g., GRE, LSAT, MCAT), and tests used for educational and vocation guidance (DAT). See A. Anastasi & S. Urbina, Psychological Testing (7th ed. 1997); R. Gregory, Aptitude Tests, in The Encyclopedia of Intelligence 110-117 (R. Sternberg, ed., 1994); R. Snow, Abilities and Aptitudes, in Encyclopedia of Intelligence 3-5 (1994).

This *design-for-prediction/selection test* battery approach employed by the ASVAB (and other differential aptitude batteries) differs significantly from the *design-for-representative sampling of abilities that represent general intelligence* in the case of general intelligence test batteries. As noted in Anastasia's classic book on psychological testing, aptitude tests (like the ASVAB) were developed *after* global IQ tests had already been established in psychological testing "to supplement the global intelligence tests—these special aptitude tests were developed particularly for use in vocational counseling and in the selection and classification of industrial and military personnel." See A. Anastasi & S. Urbina, Psychological Testing 40 (1997). Although differential aptitude tests like the ASVAB include abilities from some of the same CHC domains as tests of general intelligence, they also sample abilities from *other domains (e.g., domain-specific specialized acquired knowledge in electronics, shop, auto, etc.)*

Gc Cross-Battery Assessment (1998); L. Phelps, K.S. McGrew, S.N. Knopik, & L. Ford, The General (*g*), Broad, and Narrow CHC Stratum Characteristics of the WJ III and WISC-III Tests: A Confirmatory Cross-Battery Investigation, 20(1) School Psychology Quarterly 66-88 (2005); R.D. Roberts, G.N. Goff, F. Anjoul, P.C. Kyllonen, G. Pallier & L. Stankov, The Armed Services Vocational Batter (ASVAB) – Little More Than Acculturated Learning (Gc)!?, 12(1) Learning and Individual Differences 81-103 (2000); G. Taub, K.S. McGrew, & E. Witta, A Confirmatory Analysis of the Factor Structure and Cross-Age Invariance of the Wechsler Adult Intelligence Scale-Third Edition, 16 Psychological Assessment 85-89 (2004); R.W. Woodcock, Theoretical foundations of the WJ-R measures of Cognitive Ability, 8 Journal of Psychoeducational Assessment 231-258 (1990).

[3] The ASVAB underwent a number of revisions and changes in the late 1990's and early 2000's. The current report does not generalize to the most current versions of the ASVAB and addresses the ASVAB used prior to these later revisions and enhancements.

that are *not considered appropriate* for inclusion in a battery intended to provide a full-scale global IQ score to represent general intelligence. See e.g., L. Cronbach, The Armed Services Vocational Aptitude Battery – A Test Battery in Transition, Professional and Guidance Journal 232-237 (Jan. 1979).

The critically important difference between aptitude (ASVAB) and general intelligence tests are:

1. Individually administered IQ tests sample from the human ability domains known to represent different aspects of general intelligence (g) while differential aptitude tests (e.g., ASVAB) selectively sample from some of the known domains of general intelligence and also sample from other human ability domains not considered appropriate for inclusion in a measure of general intelligence (g).

2. The design and statistical procedures used to generate composite scores in individually administered IQ tests (e.g., Full Scale IQ; General Intellectual Ability score) and differential aptitude batteries (e.g., ASVAB AFQT score) are *markedly different in form and function.* The design and validation of IQ test batteries is to provide a global composite IQ score that is the best possible indicator of the domain of general intelligence while differential aptitude battery design and validation is focused on generating scores that are the best empirical predictors of external and circumscribed performance criteria (e.g., success in military and/or occupational training programs or settings).

Description of the ASVAB and AFQT scores

The version of the ASVAB addressed in this report includes the following paper-and-pencil tests which are administered in a group testing format:[4]

1. General Science (GS): 25 science-fact based items.
2. Arithmetic Reasoning (AR): 30 arithmetic word problems.
3. Word Knowledge (WK): 35 standard vocabulary items.
4. Paragraph Comprehension (PC): 15 (three sentence length) paragraphs that are read followed by multiple choice questions regarding each paragraph's content.
5. Numerical Operations (NO): A 10 minute speeded test were subject answers 50 (simple) number-fact items.
6. Coding Speed (CS). A 10 minute speeded test (84 items) were a word is followed by five four-digit number strings and the subject looks up the word's number code in a key consisting of 10 word-code pairs at the top of the page, and then selects the letter associated with that number code.
7. Auto and Shop Information (AI/SI): 25 questions about automobiles, shop practices, and conventional use of mechanical tools.
8. Mathematics Knowledge (MK): 25 mathematical problems.
9. Mechanical Comprehension (MC): 25 questions, normally accompanied by drawings related to general mechanical and physical principles.
10. Electrical Information (EI): 20 questions related to electrical, radio and electronics information.

Simple content task analysis of the above 10 tests reveals 3 tests that measure obvious *acquired and specialized domain-specific information and knowledge not represented on individually administered general intelligence tests* (viz., Auto and Shop Information; Mechanical Comprehension; Electrical

[4] See R.D. Roberts, G.N. Goff, F. Anjoul, P.C. Kyllonen, G. Pallier & L. Stankov, The Armed Services Vocational Batter (ASVAB) – Little More Than Acculturated Learning (Gc)!?, 12(1) Learning and Individual Differences 81-103 (2000).

Information). The *content validity* criticism has plagued the ASVAB since its inception. For example, in 1979, Lee Cronbach, a historically prominent and influential educational and psychological measurement expert, stated (in the context of the value of the ASVAB as a vocational aptitude counseling tool) that "Subtests *AI, SI, and EI being measures of experience and not talent,* have extremely limited value for counseling. To judge a person as lacking aptitude for trades on the basis of an information *test is inappropriate and damaging."* L. Cronbach, The Armed Services Vocational Aptitude Battery – A Test Battery in Transition, Professional and Guidance Journal 233 (Jan. 1979) (emphasis added). Although Cronbach's criticism may be dated, these non-intelligence subtests continue to be included in the ASVAB, most likely for the same reason articulated by Cronbach in 1979—"The information subtests are of interest to the military because they assess immediate readiness to take up certain service specialties after brief training." Id. at 235. This is *not* a scientifically accepted criterion for selection of subtests to include in test batteries intended to measure general intelligence (g).

Furthermore, although a number of the ASVAB tests share common content with subtests on some individually administered IQ tests (e.g., Arithmetic Reasoning; Word Knowledge; Numerical Operations; Coding Speed), they differ significantly from IQ subtests in that the subject *must read the items and possible answers*—thus introducing the unrelated ability of reading (aka., *construct irrelevant variance*) into the desired test score.

Relevant ASVAB Research: The ASVAB is not a measure of general intelligence

The ASVAB has been subjected to extensive internal and external validity research. See A. Anastasi & S. Urbina, Psychological Testing (1997). Much of the early ASVAB research was of questionable quality. See L. Cronbach, The Armed Services Vocational Aptitude Battery – A Test Battery in Transition, Professional and Guidance Journal 233 (Jan. 1979). According to Cronbach, "I have seen about a dozen research reports on ASVAB. Very few, I think, would have been accepted for publication by a referred journal. Research on ASVAB has not come up to the best traditions in military psychology." Id. at 236.

Of particular relevance for proper and valid interpretation of ASVAB scores is contemporary factor analysis research, which is the primary and traditional method for establishing the *"g" or general intelligence* factor characteristics of a battery of tests. According to Anastasia and Urbina, the considerable factor analytic research of the ASVAB has typically shown a "general *factor*" and four group factors that measure more specialized verbal, speed, quantitative and technical abilities or factors. See A. Anastasi & S. Urbina, Psychological Testing (1997)

It is important to note that Anastasia and Urbina do not call the ASVAB general factor a "general *intelligence*" factor--but instead, a "general *factor*." It is recognized by measurement experts in the cognitive and human ability research literature that any battery of tests that measure a diverse array of related abilities will produce a *single general factor* when subjected to factor analysis methods (the principle of *positive manifold*—the tests are all positively correlated[5]). Yet, the *mere presence of a single general factor does not necessarily indicate that it is a good indicator for general intelligence (g).*

Contemporary intelligence scholars and measurement experts now recognize that one critical source of validation for an intelligence battery (as a measure of general intelligence) requires more than performing internal statistical analysis (factor analysis) with only the specific tests within a psychological

[5] L. Humphreys, Measurement and Prediction of Intelligence, in Encyclopedia of Intelligence 694-697 (R. Sternberg, ed., 1994).

battery (aka., *within-battery* factor analysis) but, also requires *joint- or cross-battery* factor analysis studies.[6]

Collectively, factor analysis research of the ASVAB indicates that the ASVAB provides for limited measurement of the cognitive abilities from the broad ability domains from which general intelligence tests draw tests and items. Finally, and more importantly, the fact that 6 (7 if Math Knowledge is also included) of the 10 ASVAB tests require the subject to *read* the test items and answers and measure *acquired domain-specific specialized knowledge* (e.g., auto, shop, electrical and mechanical information and knowledge) *not* considered valid for the measurement of general intelligence, makes it clear that the ASVAB should *not* be considered a valid proxy or indicator of general intelligence (*g*) as measured by traditional IQ tests used to diagnose mental retardation / intellectual disability.

Furthermore, the ASVAB-based *AFQT score* is an even more narrow measure of a limited set of abilities that should *not* be equated with the general intelligence (*g*) measured by individually administered IQ tests. The AFQT is an equation-based score ([2 x Verbal score] + Arithmetic Reasoning score + Mathematical Knowledge score). It is important to note that the ASVAB AFQT score is *based on only 3 ASVAB subtests* (one that is weighted two times the others). In this regard the ASVAB AFQT score should be accorded the status of a *brief screening measure*, much like the WASI (Wechsler Abbreviated Scale of Intelligence), and *not* a score that can be interpreted as a valid indicator of general intelligence as measured by a comprehensive test of general intelligence.

Based on a review of the available published and unpublished ASVAB research literature, it is concluded that the ASVAB AFQT score should *not* be considered a valid proxy of general intellectual functioning (g). The ASVAB AFQT screening score represents a very limited set of human abilities for the specific purpose of satisfying a narrow set of military prediction/selection requirements. The ASVAB test battery and the ASVAB-derived AFQT score should *not* to be interpreted as valid measures of general intelligence.[7]

---

[6] See M.H. Daniel, Intelligence Testing: Status and Trends, 52(10) American Psychologist 1038-1045 (1997); K.S. McGrew, Analysis of the Major Intelligence Batteries According to a Proposed Comprehensive Gf–Gc Framework, in Contemporary Intellectual Assessment: Theories, Tests, and Issues 151−179 (D. P. Flanagan, J. L. Genshaft, & P. L. Harrison, eds., 1997); K.S. McGrew, The Cattell–Horn–Carroll Theory of Cognitive Abilities, in Contemporary Intellectual Assessment: Theories, Tests, and Issues 136−181 (D. P. Flanagan & P. L. Harrison, eds., 2005); K.S. McGrew & D.P. Flanagan, The Intelligence Test Desk Reference (ITDR): Gf–Gc Cross-Battery Assessment (1998); Flanagan, McGrew & Ortiz, 2000; R.D. Roberts, G.N. Goff, F. Anjoul, P.C. Kyllonen, G. Pallier & L. Stankov, The Armed Services Vocational Batter (ASVAB) – Little More Than Acculturated Learning (Gc)!?, 12(1) Learning and Individual Differences 81-103 (2000); R.D. Roberts, P.M. Markham, G. Matthews, & M. Zeidner, Assessing Intelligence: Past, Present and Future, in Handbook of Understanding and Measuring Intelligence 333-360 (O. Wilhelm and R. W. Engle, eds., 2005); R.W. Woodcock, Theoretical foundations of the WJ-R measures of Cognitive Ability, 8 Journal of Psychoeducational Assessment 231-258 (1990).

The need to include externally validated measures of cognitive abilities when attempting to determine the cognitive abilities measured by an intelligence battery has a long history. Such earlier studies were often called “factor reference” studies. See J.B. Carroll, Human Cognitive Abilities: A Survey Factor Analytic Studies (1993)

[7] Attempts to find published research that reported correlations between ASVAB scores and individually administered comprehensive IQ test scores produced no useful results. Using the PsychINFO ® data base on 5-14-10 (PsycINFO is an abstract database that provides systematic coverage of the psychological literature from the 1800s to the present; American Psychological Association-APA) and using various combinations of search terms (ASVAB and IQ; ASVAB and WAIS; ASVAB and intelligence; ASVAB and WAIS-R) produced no useful results for studies that reported correlations between ASVAB scores and scores from individually administered general intelligence tests accepted in the diagnosis of mental retardation / intellectual disability.

Interpretation of the ASVAB as a measure of general intelligence (*g*) is inconsistent with the Joint Test Standards

As prescribed in the Joint Test Standards, test users have a number of responsibilities to insure that inferences drawn from tests scores are valid. These include:

> Standard 1.4: "If a test is used in a way *that has not been validated*, it is *incumbent on the user to justify the new use*, collecting new evidence if necessary."
>
> Standard 11.15: "Test users *should be alert to potential misinterpretations of test scores and to possible unintended consequences of test use*; users should *take steps to minimize or avoid foreseeable misinterpretations and unintended negative consequences."*
>
> Standard 11.16: "Test users should *verify periodically that their interpretations of test data continue to be appropriate,* given any significant changes in their population of test takers, their modes of test administration, and *their purposes of testing*."

See Joint Test Standards at 18, 116-117 (emphasis added).

The interpretation of the ASVAB as a measure of '*g*" (general intelligence)" is inconsistent with professional adherence to Joint Test Standards 1.4, 11.15, and 11.16. A review of armed services special research reviews and reports indicate that the ASVAB was designed and validated for making statements regarding the prediction of probable success of individuals as per clearly circumscribed and specific military performance criteria. See J. Welsh, J. Kucinkas, & L. Curran, Armed Services Vocational Aptitude Battery (ASVAB): Integrative Review of Validity Studies (AFHRL-TR-90-63), Brooks, AFB, TX: U.S. Air Force Human Resources Laboratory (1990); J. Welsh, T. Watson, & M. Ree, Armed Services Vocational Aptitude Battery (ASVAB): Predicting Military Criteria from General and Specific Abilities (AFHRL-TR-90-22), Brooks, AFB, TX: U.S. Air Force Human Resources Laboratory (1990); J. Wolfe, D. Alderton, G. Larson, & J. Held, Incremental Validity of Enhanced Computer Administered Testing (ECAT), (NPRDC-TN-96-6), San Diego, CA: Navy Personnel Research and Development Center (1993). Although some of these reports do report research on "general factors" extracted from the ASVAB, these general factors are then used as to predict success in military settings, **not** to provide evidence of the ASVAB as a measure of general intelligence.

As per Joint Test Standard *1.4*, for a psychologist to claim that the ASVAB is a measure of general intelligence, it would be the psychologist's professional responsibility to provide evidence to support this claim, a claim not based on the ASVAB's program of internal technical research nor independent research (summarized above).

Summary Conclusions Regarding Interpretation of ASVAB Test Scores

The interpretation of the ASVAB as a measure of general intelligence, and particularly the interpretation of the derived 3-subtest AFQT score (without explanation of the empirical derivation of this *brief/screening* AFQT score), has the potential to mislead individuals who are unfamiliar with the ASVAB's conceptual and psychometric derivation (and contemporary ASVAB research) to inappropriately conclude

that the ASVAB measures general intelligence, the first prong in the diagnosis of mental retardation / intellectual disability. A review of the scientific research evidence indicates **that the ASVAB (and any of its derived scores) is <u>not</u> a valid proxy or indicator of *general intellectual ability*.** Attempts to interpret ASVAB scores as an indicator of *g* (general intelligence) is at variance with evidence-based standards for psychological test interpretation and has the potential to lead uninformed readers to the erroneous conclusion that an individual's ASVAB scores (if above the MR/ID range) could be considered as confirmatory evidence of a non-mental retardation / intellectual disability diagnosis. **Individual ASVAB scores *should not be used* in any form to support (or not support) a diagnosis of an individual of mental retardation / intellectual disability**.

[<u>Note</u>: A series of figures are presented in <u>Appendix A</u> that illustrate the information and concepts presented in this report]

Author information and conflict of interest disclosure

Dr. Kevin S. McGrew, Ph.D., is an Educational Psychologist with expertise and interests in applied psychometrics, intelligence theories and testing, human cognition, cognitive and non-cognitive individual difference variables impacting school learning, models of personal competence, conceptualization and measurement of adaptive behavior, measurement issues surrounding the assessment of individuals with disabilities, brain rhythm and mental timing research, and improving the use and understanding of psychological measurement and statistical information by professionals and the public. Prior to establishing IAP, Dr. McGrew was a practicing school psychologist for 14 years. McGrew received his Ph.D. in Educational Psychology (Special Education) from the University of Minnesota in 1989.

Dr. McGrew is currently Director of the *Institute for Applied Psychometrics* (IAP), a privately owned applied research organization established by McGrew. He is also the *Research Director for the Woodcock-Munoz Foundation* (WMF), Associate Director for *Measurement Learning Consultants* (MLC), and a *Visiting Professor in Educational Psychology* (School Psychology) at the University of Minnesota.

Dr. McGrew authored the current document in his role as the Director of IAP. The opinions and statements included in this report do not reflect or represent the opinions of WMF, MLC, or the University of Minnesota.

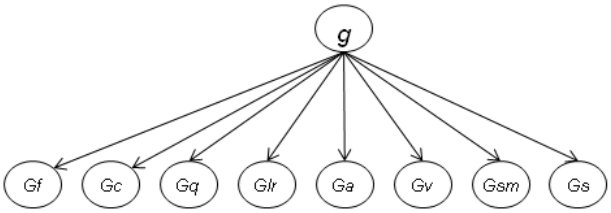
More complete professional information, including his professional resume, bio, and conflict of interest disclosures can be found at each of his three professional blogs and web page:

- www.iqscorner.com
- www.atkinsmrdeathpenalty.com
- www.ticktockbraintalk.blogspot.com
- www.iapsych.com

Appendix A

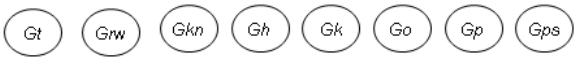
Supplementary figures for understanding why the ASVAB should not be considered as a measure of general intellectual functioning

Major cognitive ability domains sampled across the major individualized IQ batteries (Wechslers, Stanford-Binet, WJ III/BAT III) which are combined to produce general intelligence (g) full-scale global composite IQ score



What we currently know about the taxonomy of human abilities, The primary CHC broad ability domains (McGrew, 2009)

Other human ability domains identified in the research literature but not typically included in IQ batteries, although some abilities may be included in special differential aptitude batteries



CHC Broad (Stratum II) Ability Domains

Gf	Fluid reasoning
Gc	Comprehension-knowledge
Gsm	Short-term memory
Gv	Visual processing
Ga	Auditory processing
Glr	Long-term storage and retrieval
Gs	Cognitive processing speed
Gt	Decision and reaction speed
Grw	Reading and writing
Gq	Quantitative knowledge
Gkn	General (domain-specific) knowledge
Gh	Tactile abilities
Gk	Kinesthetic abilities
Go	Olfactory abilities
Gp	Psychomotor abilities
Gps	Psychomotor speed

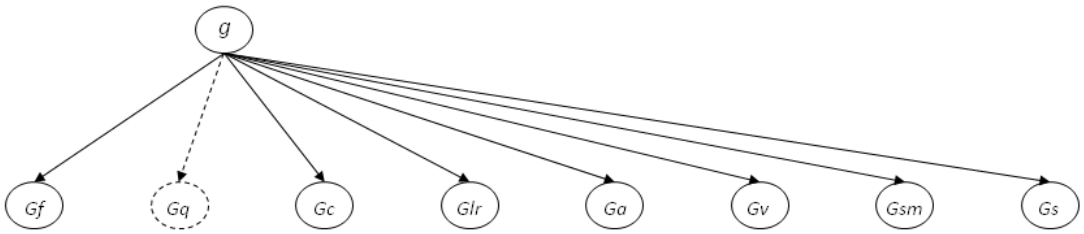
(see Table 1 for definitions)

Complete CHC model and description of abilities can be found in:

McGrew, K. (2009). Editorial: CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research, *Intelligence, 37,* 1-10.

Stratum III *g*-factor is offset to left as per Carroll (1993) to reflect degree to which broad (stratum II) *Gf-Gc* abilities are correlated with *g* in the extant literature.

Portions of Cattell-Horn-Carroll (CHC) theory of cognitive abilities measured by some or all major contemporary intelligence batteries



70+ narrow (stratum I) abilities have been identified but are not included in figure for readability purposes

Dashed *Gq* broad ability arrow and oval designates that math achievement abilities are typically found in achievement tests, but have been shown to be measured by some tests in some cognitive/IQ batteries (e.g., Wechsler Arithmetic subtest)

CHC Broad (Stratum II) ability domains included across cognitive test batteries

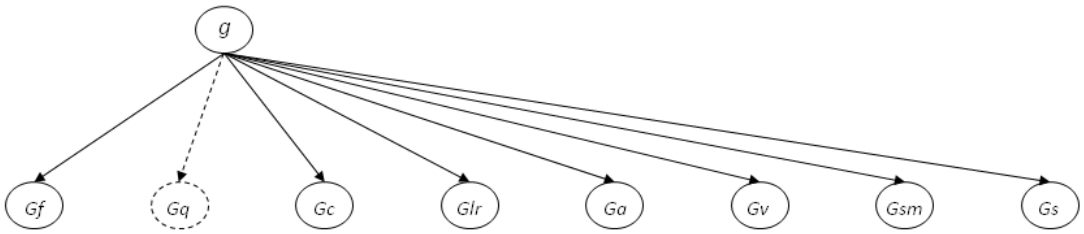
Gf	Fluid reasoning
Gc	Comprehension-knowledge
Gsm	Short-term memory
Gv	Visual processing
Ga	Auditory processing
Glr	Long-term storage and retrieval
Gs	Cognitive processing speed
Gq	Quantitative knowledge

Complete CHC model and description of abilities can be found in:

McGrew, K. (2009). Editorial: CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research, *Intelligence, 37*, 1-10.

Stratum III *g*-factor is offset to left as per Carroll (1993) to reflect degree to which broad (stratum II) *Gf-Gc* abilities are correlated with *g* in the extant literature.

Portions of Cattell-Horn-Carroll (CHC) theory of cognitive abilities measured by some or all major contemporary intelligence batteries



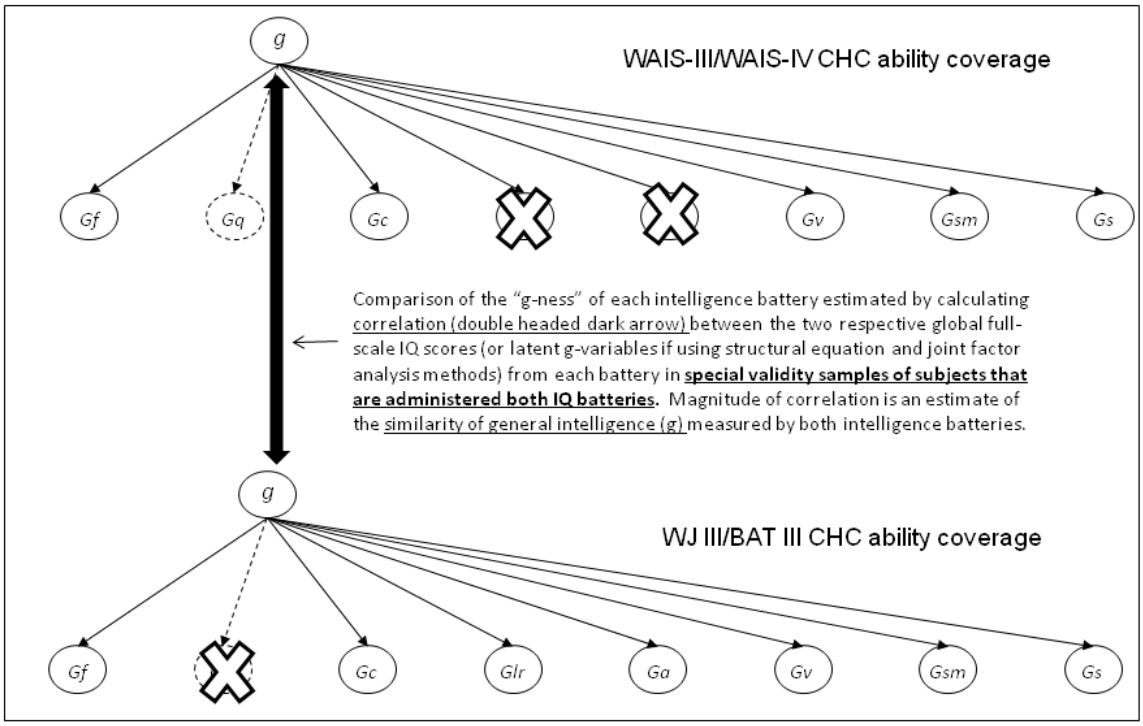
70+ narrow (stratum I) abilities have been identified but are not included in figure for readability purposes

Dashed *Gq* broad ability arrow and oval designates that math achievement abilities are typically found in achievement tests, but have been shown to be measured by some tests in some cognitive/IQ batteries (e.g., Wechsler Arithmetic subtest)

A primary design goal of a comprehensive IQ battery is to provide a global full-scale IQ index that provides the best **possible estimate of "general intelligence (g)"** by sampling from the major domains of human intelligence (viz., CHC broad cognitive ability domains)

IQ batteries differ in the number of CHC domains sampled and emphasized in the calculation of each batteries respective global full-scale IQ score

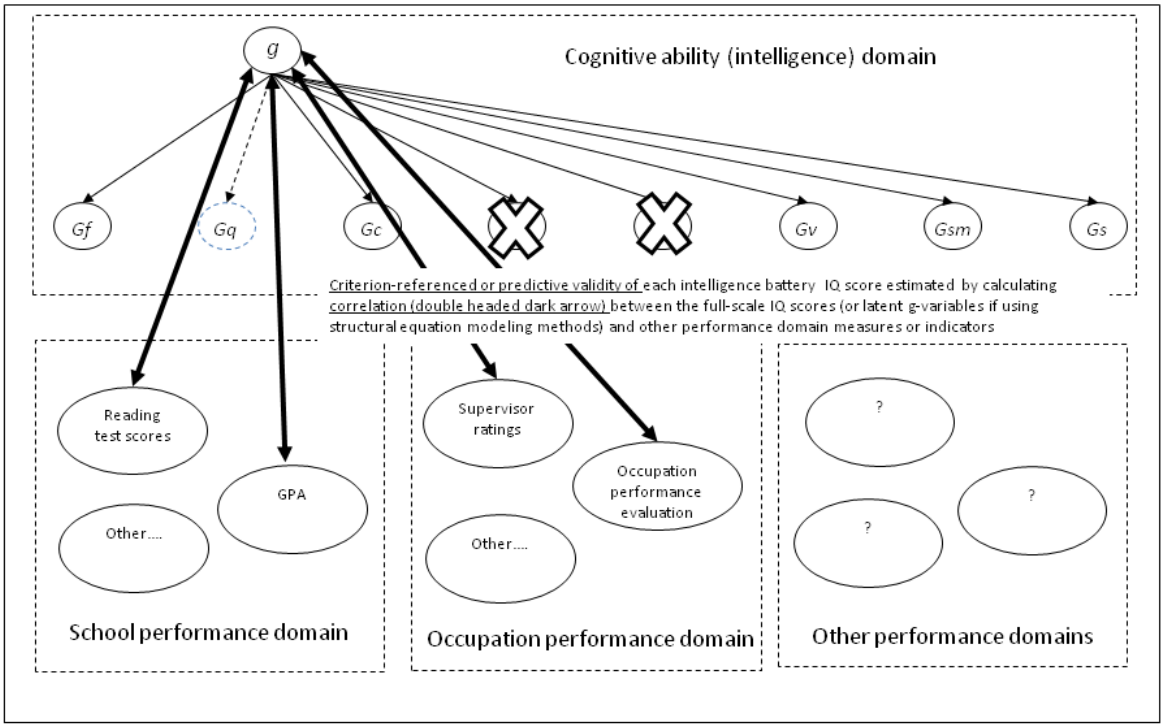
How to validate the "g-ness" of intelligence batteries: One primary method—concurrent validity correlations with other validated measures of general intelligence



Rectangle represents single validity sample of subjects who are administered both tests

Ability domain not measured by IQ battery

How to validate the "g-ness" of intelligence batteries: One secondary method—criterion or predictive validity – goal is have IQ score predict a wide array of human behavior outcomes or performances

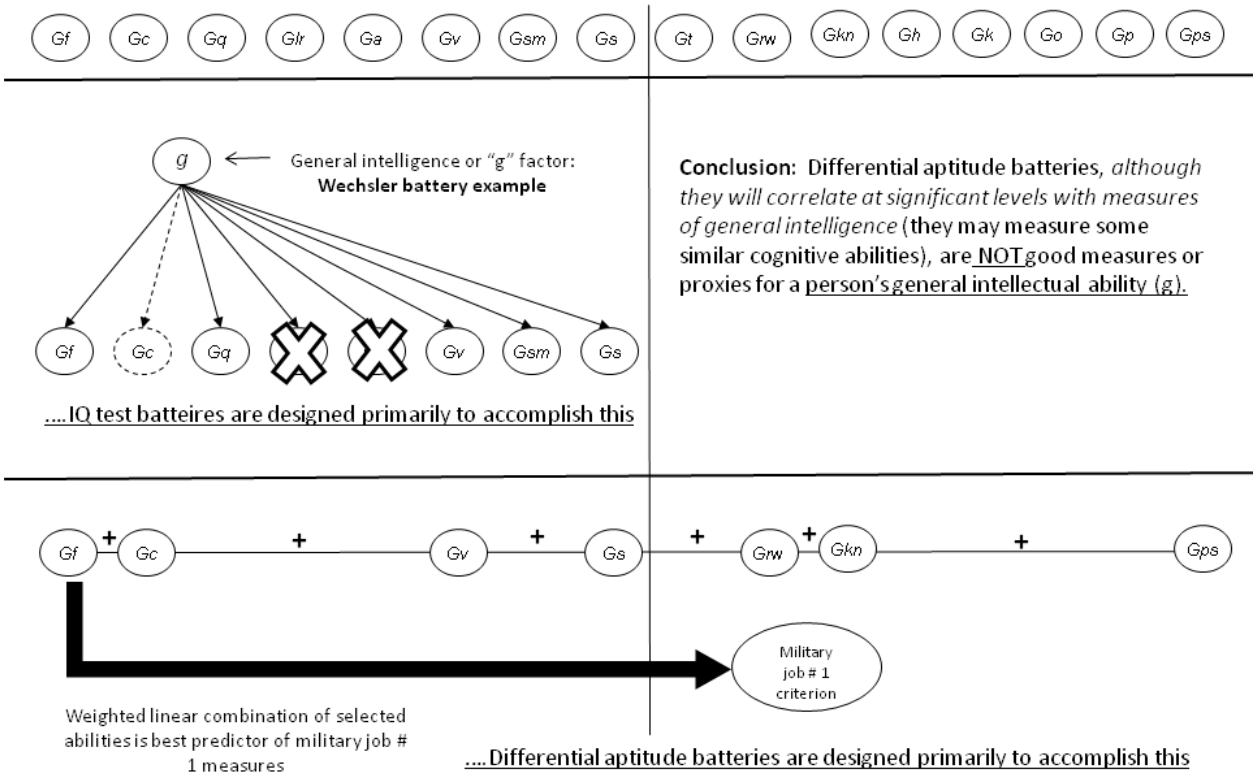


Solid rectangle represents single validity sample of subjects who are administered IQ battery plus other criterion measures

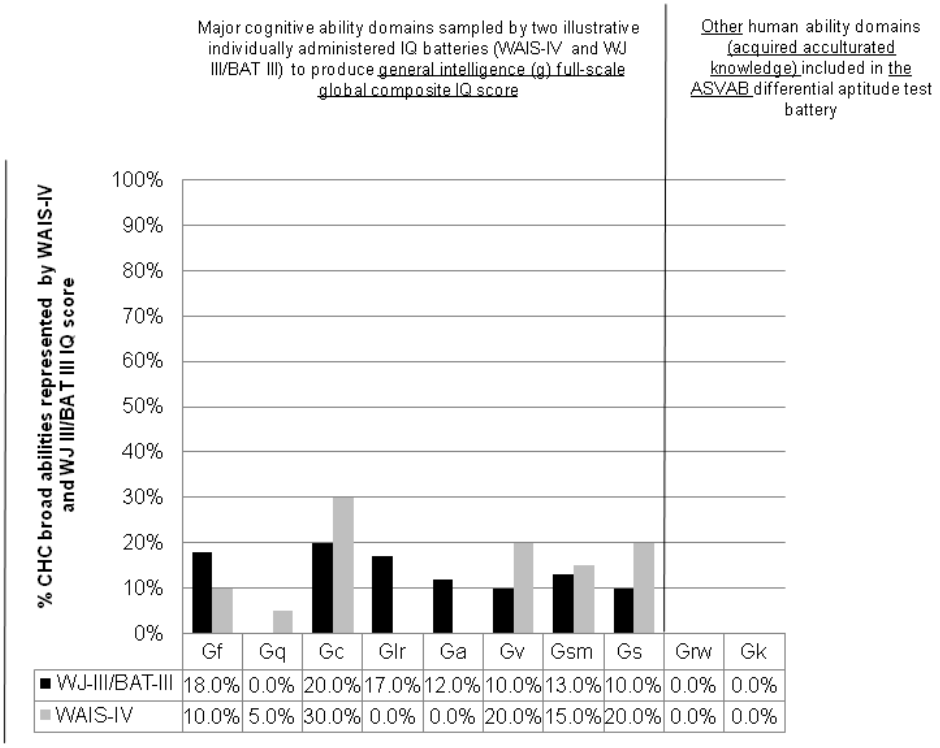
Dashed rectangles represents different human performance domains. Subjects in special validity sample(s) are administered IQ test together with other criterion performance measures to establish different types of validity evidence

Primary (critical) difference in purpose-based design of general intelligence (IQ) and differential aptitude test batteries

Given the current "working" taxonomy of human abilities (primary CHC broad ability domains)....

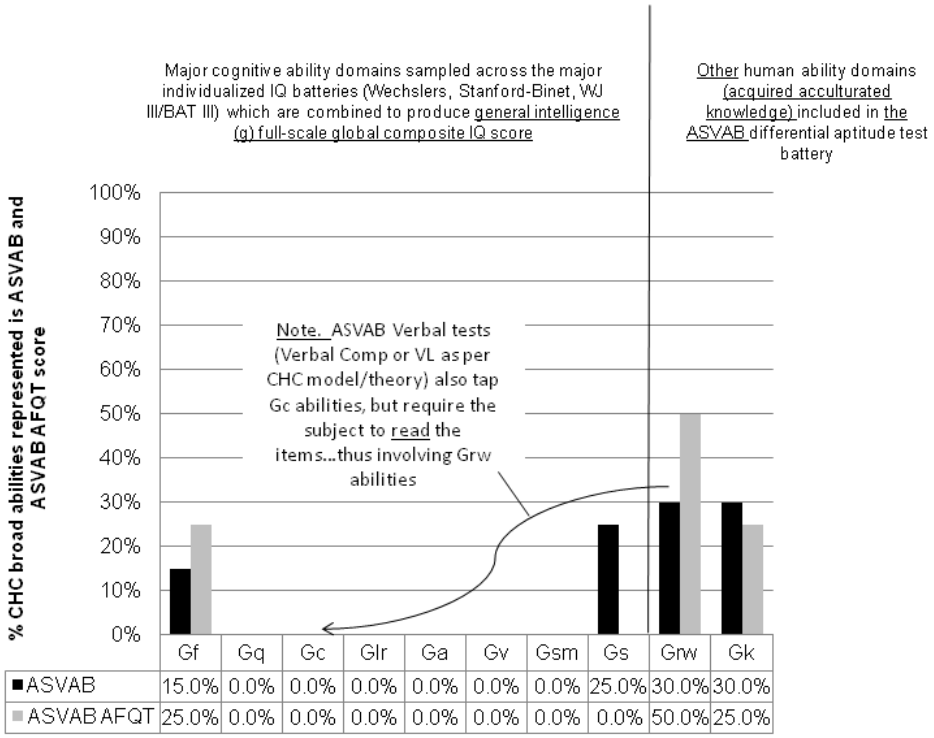


Proportional **CHC broad ability** coverage of two illustrative IQ test batteries (WAIS-IV & WJ III/BAT III)



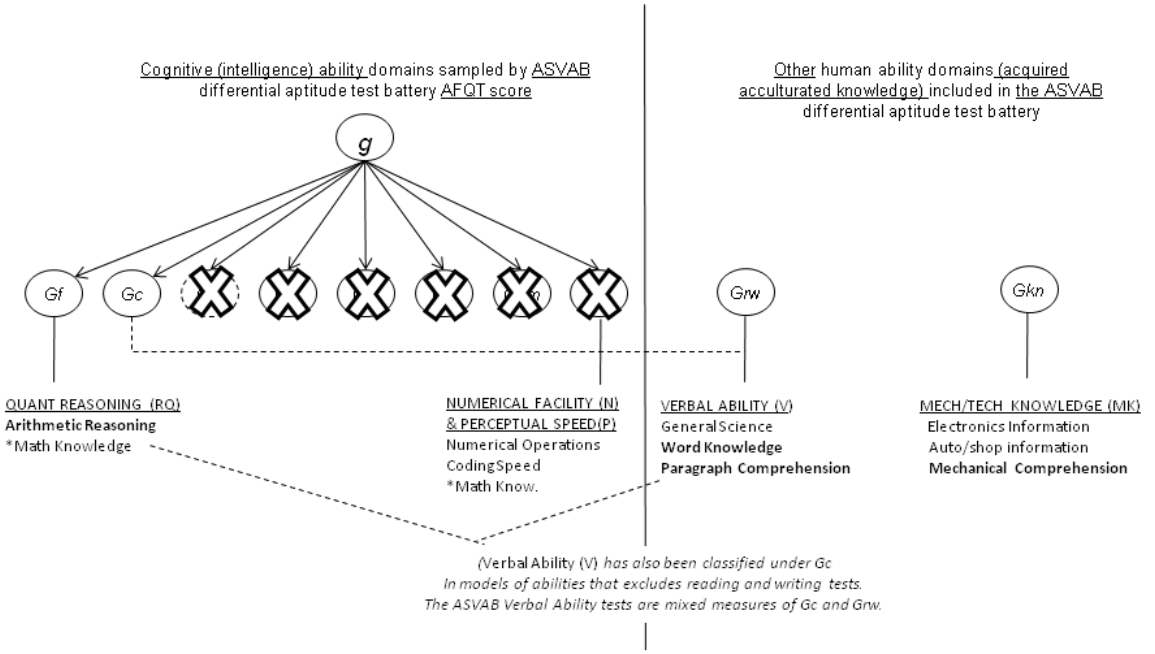
	Gf	Gq	Gc	Glr	Ga	Gv	Gsm	Gs	Grw	Gk
WJ-III/BAT-III	18.0%	0.0%	20.0%	17.0%	12.0%	10.0%	13.0%	10.0%	0.0%	0.0%
WAIS-IV	10.0%	5.0%	30.0%	0.0%	0.0%	20.0%	15.0%	20.0%	0.0%	0.0%

Proportional **CHC broad ability** coverage of ASVAB and ASVAB-derived AFQT score



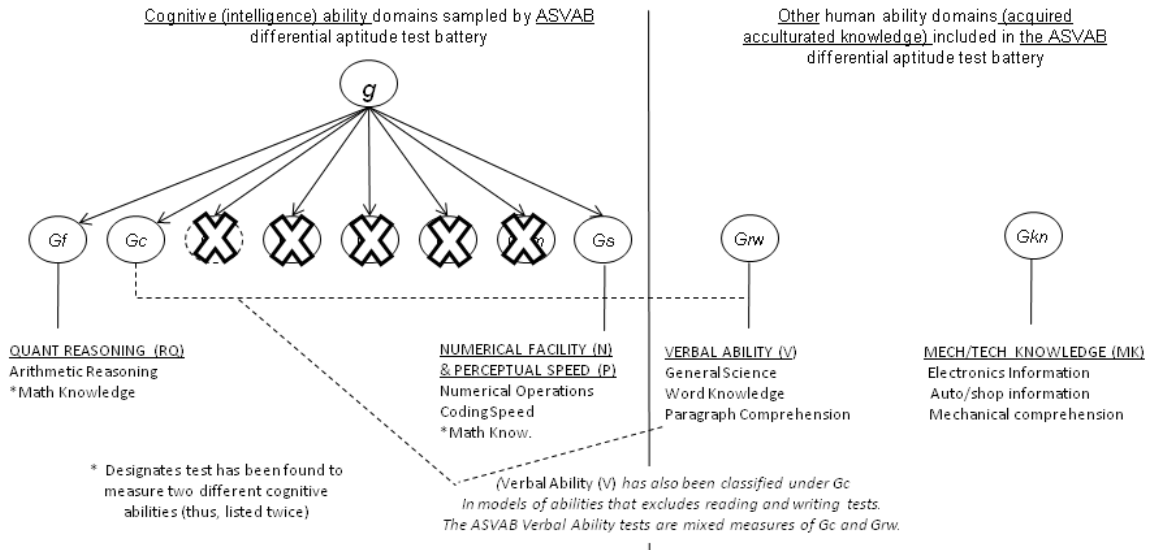
	Gf	Gq	Gc	Glr	Ga	Gv	Gsm	Gs	Grw	Gk
ASVAB	15.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	25.0%	30.0%	30.0%
ASVAB AFQT	25.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	50.0%	25.0%

ASVAB AFQT score is based on following equation: [2 x (Verbal)] + Arithmetic Reasoning + Math Knowledge
(Verbal = Word Knowledge + Paragraph Comprehension)



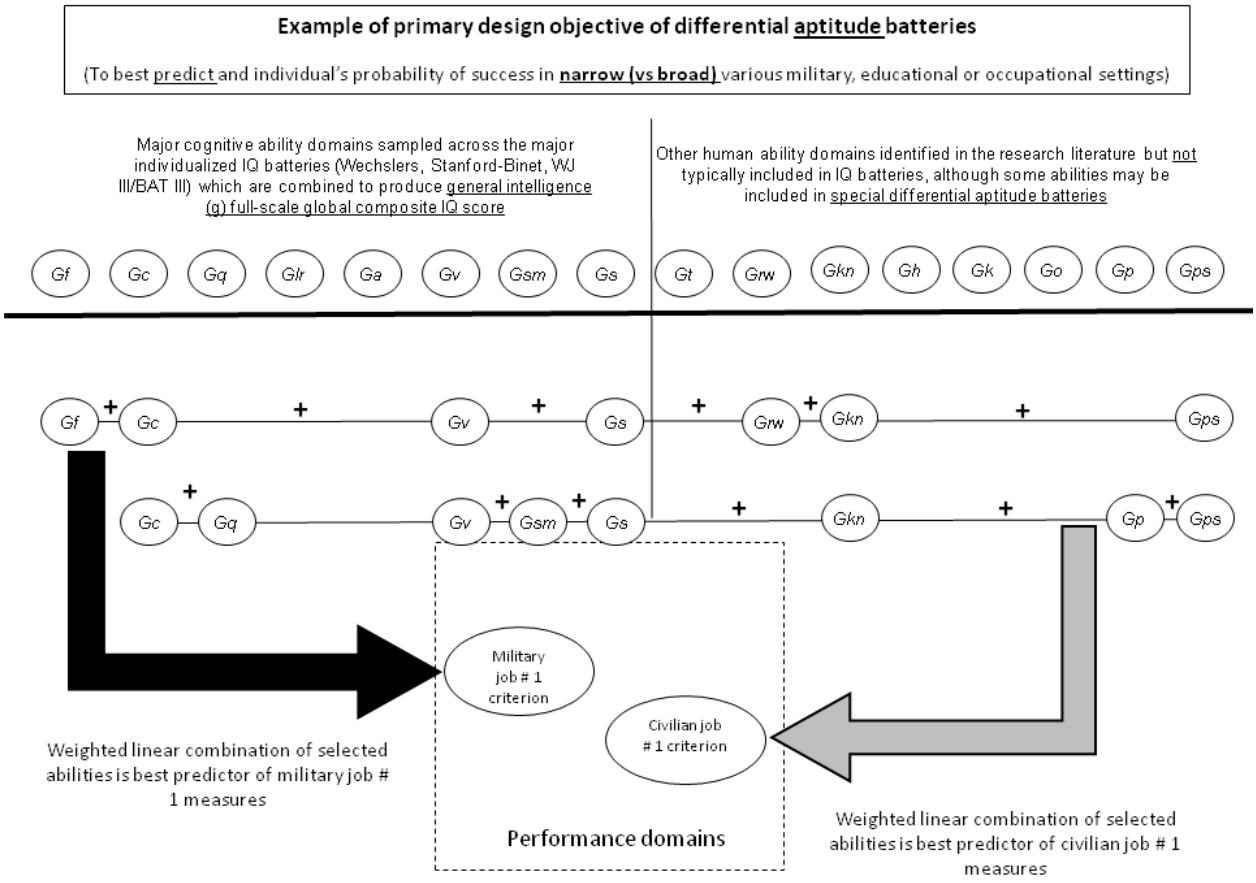
Conclusion: ASVAB AFQT score is a narrow composite of abilities and is not representative or comparable to a broad-based global composite full-scale IQ score

CHC ability content coverage of ASVAB in independent factor analysis research



Summary based on contemporary CHC intelligence model-based (McGrew, 2009) syntehsis of research studies by:

- Carroll, J. (1993). *Human cognitive abilities: A survey of factor analytic studies.* New York: Cambridge University Press.
- Gustafsson, J.-E., & Muthén, B. (1994). *The nature of the general factor in hierarchical models of the structure of cognitive abilities: Alternative models tested on data from regular and experimental military enlistment tests.* Technical Report: University of Göteborg, Sweden.
- Roberts, R., Goff, G., Anjoul, Kyllonen, P., Pallier, & Stankov, L. (2000). The Armed Services Vocational Aptitude Battery (ASVAB): Little more than acculturated learning (Gc) ! ? *Learning and Individual Differences, 12,* 81-103.
- Wolfe, J., Alderton, D., Larson, G., & Held, (1993). *Incremental validity of enhanced computer administered testing (ECAT).* Manuscript. Navy Personnel Research and Development Center, San Diego, California


Example of primary design objective of differential aptitude batteries
(To best predict and individual's probability of success in narrow (vs broad) various military, educational or occupational settings)
Major cognitive ability domains sampled across the major individualized IQ batteries (Wechslers, Stanford-Binet, WJ III/BAT III) which are combined to produce general intelligence (g) full-scale global composite IQ score
Other human ability domains identified in the research literature but not typically included in IQ batteries, although some abilities may be included in special differential aptitude batteries
Gf
Gc
Gq
Glr
Ga
Gv
Gsm
Gs
Gt
Grw
Gkn
Gh
Gk
Go
Gp
Gps
Military job # 1 criterion
Civilian job # 1 criterion
Performance domains
Weighted linear combination of selected abilities is best predictor of military job # 1 measures
Weighted linear combination of selected abilities is best predictor of civilian job # 1 measures