The Use of Calculators on the ASVAB

Executive Summary

With the continual interest in calculator usage during ASVAB administrations, OPA seeks to make every effort to help stakeholders understand the rationale and theory behind the ASVAB. We offer this document to provide information, context, and clarity to our recruiters and stakeholders on the history of and decisions regarding calculator use.

After careful review of the history and intended use of the ASVAB, and an extensive literature review of calculator use on various large-scale math assessments, OPA re-evaluated the use of calculators and re-examined the advantages and disadvantages of allowing calculators on the current ASVAB. OPA concludes the following:

- that allowing calculator use would not increase the number of eligible recruits due to the test scaling and equating requirements;
- that while allowing calculators may give the appearance that the ASVAB is keeping up with the latest trends and technologies and would better align with college entrance exams, the ASVAB and college entrance exams serve different purposes. Moreover, it lacks the technical and psychometric merit necessary for such a change, and in fact, the reliability and validity of the ASVAB scores would be at risk; and
- the costs of allowing calculators outweigh any possible benefits given the complexity of necessary re-standardization processes.

The process of introducing substantive changes to the ASVAB is rigorous and complex. Given the above, OPA does not recommend that Accession Policy change the current policy on the use of calculators.

Current Policy

One of the most commonly asked questions test takers of the ASVAB pose is whether calculator use is allowed on the ASVAB. The answer based on current policy is no.

Why Don't We Allow Calculators?

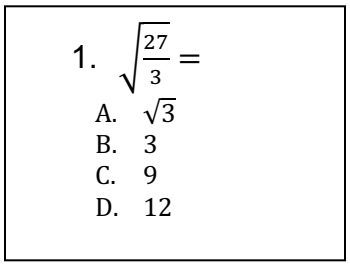
First introduced in 1968, the ASVAB has been a pioneer in aptitude testing. Without a doubt, the ASVAB is one of the most well-respected and researched tests in modern testing history. Such an achievement is possible only by consistently adhering to the highest standards in testing and measurement to ensure the validity and reliability of the test scores.

Administered annually to more than one million military applicants (ASVAB ETP) and high school students (ASVAB CEP), the ASVAB has been a vital component of the United States Armed Forces selection and classification system. Examinees receive ASVAB standard scores and norm-based scores

on various subtests and composites. These ASVAB scores are used to determine enlistment eligibility, assign applicants to military occupational specialties, and aid students in career exploration.

When the ASVAB was introduced in the 1960s, calculators were rare, and the ability to conduct hand calculations in the field was required by many Services. The ASVAB items have been written based on the test specifications that examinees do not need calculators to solve the problems. Today, the requirement for hand calculations in the field is still true for some Services. Over the years, OPA (formerly Defense Manpower Data Center or DMDC) has carefully evaluated the need to allow calculators when taking the ASVAB and has decided the current calculator policy is still valid for ASVAB's various score-use purposes.

Let's begin by looking at one sample math question on the ASVAB (right panel). To answer this question without using a calculator, examinees must be able to first solve a fraction (perform division: 27/3), then find the square root of a number (9). If an examinee makes a mistake during any of these steps, he or she could choose one of the distractors instead of the correct answer (B). With a calculator, however, the correct answer could be obtained readily without following a correct process or knowing basic arithmetic concepts and mathematical operations.



This item is intended to assess examinees' ability to correctly carry out basic mathematical calculations (in this case, division and square root of a number) without using a calculator. It requires knowledge of arithmetic concepts, mathematical operations, and following the correct process; there are many similar items in the Arithmetic Reasoning (AR) and Mathematics Knowledge (MK) tests, which are the two tests in the Math domain measured by the ASVAB. These items are developed by content experts and measurement specialists to assess the abilities that are required by military service members for maintaining a highly able and trained fighting force. These test items intend to measure an examinee's math skills without a calculator because some individuals who serve in specific occupational specialties may not have ready access to calculators when quick decisions must be made that require problem-solving skills. In other words, a person's inability to solve basic math problems without the aid of a calculator could potentially have a significant impact on any number of scenarios—scenarios not as predictable as when taking a test in a classroom setting.

Should We Allow Calculators?

Whether we should allow calculators on the ASVAB is a logical question, given the prominent role technology plays in the secondary and post-secondary education sectors. However, we must also consider the potential impact on the ASVAB scores if calculators were to be permitted on current ASVAB tests with existing items that were not originally developed with calculator usage in mind.

The ASVAB is a standardized assessment, i.e., the ASVAB score scales and testing time are established under a set of fixed test directions, testing conditions, and scoring rules. The current ASVAB was standardized in 1997 through a carefully designed and rigorous data-collection and psychometric data analysis process (Segall, 2004). The 1997 standardization assured the continuance of the ASVAB score scale and established normative scores based on the 1997 Profile of American Youth study. As the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME,

2014; hereafter *Standards*) stated, "Standardization helps to ensure that all test takers have the same opportunity to demonstrate their competencies," and "The usefulness and interpretability of test scores require that a test be administered and scored according to the test developer's instructions and the conditions under which the test has been standardized. . . . The importance of adherence to appropriate standardization of administration procedures increase with the stakes of the test" (AERA, APA, & NCME, 2014). One of the ASVAB standardization conditions is that calculators are not allowed while taking the tests.

Allowing calculator use with current ASVAB tests would put the reliability and validity of the test scores at risk. As the *Standards* point out, "Deviations from standardized procedures may be inappropriate because they change the construct being measured, compromise the comparability of scores or use of norms, and/or unfairly advantage some individuals" (AERA, APA, & NCME, 2014). As demonstrated earlier with the sample math question, calculator use might affect or even alter what is intended to be measured by the item and would consequently impact the psychometric quality of the test. In fact, published empirical studies have shown that calculator use could disadvantage the less able examinees and benefit the more able ones (Attali, 2014; Scheuneman, Camara, Cascallar, Wendler, & Lawrence, 2002). An extensive literature review recently conducted by OPA on calculator use in various math tests also provides impact evidence on the following aspects of a test, which will negatively affect and threaten the validity of test scores.

Test Construct

Allowing calculator use on a test that is developed to be taken without calculators can add construct-irrelevant variance to test scores and subsequently change the construct that the test has been designed to measure. Such impact on test construct will lead to scores that differ in meaning from those of the test taken without calculators.

Item Characteristics and Statistics

The impact of calculator use on test items differs across item types. While some types of items may become easier, there are also items that may become more difficult (Attali, 2014). The differential impact on item difficulty and other characteristics will affect examinees' final scores and consequently make the validity of the existing score scale questionable. This could also degrade the accuracy of test scores in computerized adaptive testing since item selection matches question difficulty to the test-taker proficiency level.

Response Time

Calculator use tends to increase the time students need to complete a test. The more frequently a student uses a calculator, the less likely the student is to finish the test on time (Attali, 2014; Scheuneman et al., 2002). If an examinee is running out of time, he or she may not be able to provide answers to all the items or may simply rush through the test by randomly guessing the items toward the end. Either case will have a negative impact on the examinee's score. In computerized administrations of the ASVAB (CAT), examinees are penalized on the number of items they don't finish.

Test Fairness

High-ability students tend to use calculators differently from those with lower abilities. Differences in calculator usage (e.g., how often) also exist among different gender and ethnic groups, which is likely due to their varied exposure/familiarity to calculators (Attali, 2014; Scheuneman et al., 2002). The differential usage of calculators raises serious concerns about test fairness and the validity of the test score interpretations because "fairness is a fundamental validity issue" for a test (AERA, APA, & NCME, 2014).

Due to the various potential threats to the validity of the ASVAB scores, the current calculator policy on the ASVAB remains the same as the policy under which the items were developed and the tests were standardized. That is, calculators were not and are not allowed when taking the ASVAB tests.

What's Next?

Over the years, the calculator use policy on the ASVAB has been discussed and evaluated. At a July 15, 2009, meeting, OPA (then DMDC) presented various issues related to the use of calculators on the ASVAB to the technical committee for the Manpower Accession Policy Working Group (MAPWG). After a careful evaluation of all the pros and cons, the committee supported OPA's recommendations to keep the no-calculator-use policy and consider the use of calculators only after the paper-and-pencil ASVAB test administration has been eliminated (i.e., when providing standardized on-screen calculators for all examinees would become a possibility). Today, the ASVAB ETP is nearly 100% CAT, but the ASVAB CEP is still largely a paper-and-pencil test, which prohibits the introduction of a standardized calculator across all ASVAB administrations. Inclusion of a standardized on-screen calculator following paper-and-pencil elimination will be critical to establishing new standardized conditions of the assessment.

We must emphasize that any changes to the previously established standardized procedures will require re-standardization of the test under the new testing conditions. The re-standardization will address all the psychometric concerns listed in the previous section and ensure the continuance of the ASVAB score scale so military policy planners can continually "compare the cognitive level of today's force with forces of years past, to set target qualification levels, and to anticipate future trends in military needs and civilian supply" (Segall, 2004). Further, it is important to note that after re-standardization, any anticipated gains from calculator use (in the event there is a positive calcuator effect) would be statistically eliminated to ensure the continuity of the ASVAB score scale so that scores on the ASVAB with and without calculators are interchangable. In other words, if an applicant does not qualify by taking the ASVAB without a calculator, he or she will still not qualify when taking the ASVAB with a calculator.

If a future decision is made to introduce calculator use into the ASVAB, then the re-standardization process will be similar to the 1997 studies and must include the following essential tasks.

Psychometric Tasks Required for Calculator Implementation

- Develop specifications for new item development. These specifications would guide item writing under the new calculator policy in which calculator use would be allowed on the ASVAB.

- Review existing items. Existing items would need to be carefully reviewed for appropriateness given they were originally developed to assess examinees without being given access to a calculator. The new item specifications should also provide guidelines for reviewing existing items.

- Item field testing. Items would need to be field-tested with calculator usage before items are selected for operational use.

- Test scaling and equating. Sufficient data for the new test, either with newly developed items or existing items, would need to be collected to conduct scaling and equating studies so that scale scores on the new test could be placed on the existing ASVAB score scale.

- Norms development. The test would need to be re-normed. A normative population would need to be identified and a representative sample would need to be recruited to take the new test.

- Evaluate and establish new testing time. Calculator use would likely lead to the total testing time being increased. Research studies would need to be conducted on both the subtests and total time limits to ensure that sufficient time is allocated for students. In addition, practice time to allow students to familiarize themselves with the use of on-screen calculators should be taken into account, which would also add to the total testing time.

The psychometric tasks noted above (i.e., the re-standardization process) are complex and would require significant resources and time. The re-standardization process would also remove any possible (positive or negative) calculator effect on examinees' ASVAB scores and would not lead to any increase in the acceptance rate. The perceived benefit on recruitment by allowing calculators on the ASVAB, therefore, is questionable.

Operational Tasks Required for Calculator Implementation

- Implement on-screen calculator on the ASVAB delivery platform. The functionalities of the calculator should be carefully defined. The implementation needs to be tested by a focus group of potential examinees. Key issues to be addressed by the focus group study include 1) user experience: the on-screen calculator should be easy to use for all examinees; 2) potential examinee distraction: the on-screen calculator should not detract from or interfere with the test-taking experience; 3) the comparability of the on-screen calculator across different allowed devices (e.g., laptops, desktops).

- Develop a practice test with calculator use allowed or training on calculator usage. A fixed-form practice test would need to be provided so examinees can familiarize themselves with the on-screen calculator before taking the test. The practice items should include a small set of items that represent all the item types that allow calculator usage. The practice test could be provided before the operational test administration, or instruction/training could be provided on the usage of on-screen calculators to ensure test fairness.

In addition, further research studies would need to be considered in order to address the following concerns:

- The ability to conduct hand calculations in the field is still required by some Services. Allowing calculators use on the ASVAB would prevent the assessment of the hand-calculation ability. One option would be to provide two sections in the Math test: one section would allow calculator usage and the other would not (similar to the SAT, NAEP, and some state assessments).

- The unwanted consequences of allowing an on-screen calculator on the ASVAB needs to be evaluated. For example, the use of calculators could increase adverse impacts, as individuals with poorer math skills may be less adept at using calculators as some of the published studies have shown. Such a study would be critical for establishing the consequential validity of the ASVAB.

Finally, tests administered on the ASVAB platform must meet a fixed set of criteria prior to implementation. Once the psychometric and operational tasks are completed, the MAPWG technical representatives would need to evaluate whether the tests with on-screen calculators are of sufficient quality to be administered on the ASVAB platform. This evaluation process must be followed before calculator use can be implemented on the ASVAB.

Conclusion

Changes to any high-stakes assessment require careful and thorough evaluation. Given the psychometric, operational, and additional complexities around calculator use, it is clear that a possible calculator use policy change on the ASVAB is likely to take multiple steps, extensive resources, and most importantly, continuous collaborative effort and communication across various stakeholders. OPA is committed to ensuring that the ASVAB measures relevant skills and identifies examinees with the skills needed to be successful for the Services, military specialists, and their future careers. Adhering to the highest industry standards (i.e., the *Standards*) and guarding against any potential threat to the validity and reliability of the ASVAB will continue to be OPA's primary focus.

Given the various potential threats to the validity of ASVAB scores by allowing calculators on the current ASVAB, and the high cost of re-standardization of the ASVAB with calculators, OPA concludes that the costs outweigh the perceived benefits of allowing calculators on the current ASVAB. Therefore, OPA does not recommend that Accession Policy change the current policy on the use of calculators.

References

American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME), Joint Committee on Standards for Educational and Psychological Testing (U.S.) (2014). *Standards for educational and psychological testing*. Washington, DC.

Attali, Y. (2014). *Calculator use on the GRE revised general test quantitative reasoning measure*. GRE Board Research Report No. 14-02 and ETS Research Report Series No. RR-14-25. Educational Testing Service, NJ.

Scheuneman, J.D., Camara, W.J., Cascallar, A.S., Wendler, C. & Lawrence, I. (2002). Calculator Access, Use, and Type in Relation to Performance on the SAT I: Reasoning Test in Mathematics. *Applied Measurement in Education, 15*, 95–112.

Segall, D. O. (2004). *Development and evaluation of the 1997 ASVAB score scale*. (Technical Report No. 2004-002). Seaside, CA: Defense Manpower Data Center.