OFFICE OF PEOPLE ANALYTICS
Data-Driven Solutions for Decision Makers

October 2019 | EXECUTIVE NOTE | 2025-193

Why There Is No Spanish Version of the ASVAB

Executive Summary

The Office of People Analytics (OPA) typically receives inquiries each year regarding the lack of availability of a Spanish-language Armed Services Vocational Aptitude Battery (ASVAB). In response, OPA makes every effort to help stakeholders understand the rationale and theory behind an English-only ASVAB. We offer this executive note/information paper to provide additional context and clarity on the history of, and decisions regarding, an English-only ASVAB.

Given the continued interest by the Services in a Spanish-language ASVAB, OPA re-examined the potential advantages and disadvantages of developing a Spanish version of the ASVAB. After careful review of the ASVAB's history and intended use, an extensive literature review on test translation and adaptation (also known as transadaptation), as well as key statistics about English proficiency among the U.S.-born Hispanic or Latino population in the nation and in the military, OPA concludes the following:

- The population that would benefit from a Spanish-language ASVAB is likely small, given that language is expected to be a barrier for only about one in ten U.S.-born Hispanics or Latinos, and U.S. births serve as the primary source of their population growth. Additionally, the percentage of Hispanics or Latinos in the United States Armed Forces is close to the numbers in the civilian labor force: 24.0% for the Armed Forces vs. 22.6% for civilians (including Puerto Rico and other U.S. territories and possessions) (Department of Defense for Personnel and Readiness, 2017). Whether a Spanish-version ASVAB would bring in more applicants/recruits is not clear.

- Although a Spanish version of the ASVAB could allow English-learning applicants to show general aptitude in the domains measured by the ASVAB, it would not provide evidence of the applicants' ability to perform tasks in English, as would be required for training and occupational success in the United States Armed Services.

- An extremely rigorous, complex, and lengthy process would be required to develop a transadapted ASVAB. This process would need to be followed every time a new version of the ASVAB was developed.

- It would be exceedingly difficult to develop a Spanish version that is psychometrically parallel to and with the same validity and reliability as the English-language ASVAB, which could result in legal challenges to the ASVAB testing program if a Spanish version were to be developed and officially administered in conjunction with the English version.

- Developing and administering a Spanish-language ASVAB would require establishing a policy about whether to develop and administer ASVAB for alternate languages.

Given the above conclusions, OPA is confident that the costs of developing a Spanish version of the ASVAB outweigh any perceived benefits. OPA's recommendation to the Accession Policy Directorate

of the Office of the Under Secretary for Personnel and Readiness is to maintain the current policy of offering an English-only version of the ASVAB.

Current Policy

One of the most common questions raised by ASVAB stakeholders is whether an applicant can take the ASVAB in Spanish. The answer based on current policy is no. Transadaptations in languages other than English are not developed.

Why isn't there a Spanish version of the ASVAB?

Each year, the ASVAB is administered to more than one million military applicants and high school students (via the ASVAB Enlistment Testing Program and the ASVAB Career Exploration Program) and is a "multiple-aptitude battery that measures developed abilities and helps predict future academic and occupational success in the military" (http://official-asvab.com). It is a vital component of the United States Armed Forces selection and classification system. The battery is administered as a computerized adaptive test (CAT-ASVAB) or as a paper-and-pencil test (P&P-ASVAB). 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 high school students in career exploration.

There are many reasons the ASVAB is given only in English. First, all military members must be able to speak and understand English in order to successfully perform necessary assigned tasks. In the late 1970s, it became evident that the Services needed a way to assess literacy because a significant percentage of recruits, particulary in the Army, had difficulty reading instructions (in English) in their training courses. As a matter of policy, the Paragraph Comprehension subtest was added to the ASVAB in an effort to increase the reading (literacy) requirements in the Armed Forces Qualification Test (AFQT). It was later determined via research studies that the Paragraph Comprehension subtest demonstrated predictive validity (Maier & Sims, 1986). Standard 3.13 of the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014; hereafter, *Standards*) states: "A test should be administered in the language that is most relevant and appropriate to the test purpose." Strong validity evidence shows that the ASVAB scores are very good at predicting an applicant's subsequent military training performance, attrition, and on-the-job performance (Welsh, Kucinkas, & Curran, 1990; Campbell & Knapp, 2001).

Secondly, although a Spanish version of the ASVAB might allow applicants who are English-language learners to show their general aptitude in the domains measured by the ASVAB subtests, it would not provide sufficient evidence of their ability to perform the same tasks if administered in English nor provide evidence of their ability to perform in subsequent trainings conducted in English.

Thirdly, English-language learners often vary in proficiency and literacy levels in their native languages. If the ASVAB were to be translated into Spanish, the different levels of Spanish proficiency would need to be considered. Thus, it would be very difficult, if not impossible, to have a Spanish version of the ASVAB with the same validity and fairness across the Spanish-speaking population (once defined).

Finally, it is unlikely that Spanish and English versions of the ASVAB would measure exactly the same constructs (Cohen, Gafni, & Hanani, 2007; Hambleton, 2005). Scores would likely not have the same meaning across the two versions, and different evaluation and classification standards would need to be developed and used in the two versions. Under such conditions, it would be difficult to ensure that all applicants are held to the same standards.

It is also relevant to note that according to the American Community Survey (2016) conducted by the Pew Research Center, since 2000, the primary source of Hispanic or Latino population growth has been U.S. births, and consequently, English proficiency—especially among the youth—has been rising. English proficiency (speaks only English or speaks English very well) among Hispanics or Latinos who are 18 to 33 was 76% in 2014, up from 59% in 2000. More recently, 88% aged 5 to 17 were English proficient in 2014, up from 73% in 2000. Among U.S.-born Hispanics or Latinos, English proficiency rates are even higher. In 2015, the Pew Research Center reported that 89% spoke English proficiently in 2013, up from 72% in 1980 (Krogstad, Stepler, & Lopez, 2015). At the same time, only 34% of foreign-born Hispanics or Latinos were proficient in English. For people of Puerto Rican origin, 83% of mainland-born Puerto Ricans were English proficient in 2014 compared to 59% who were island-born (Lopez and Patten, 2015). However, given the population growth and English proficiency trends noted, the population that would benefit from a Spanish-language ASVAB is likely small and getting smaller every year.

Psychometric Challenges of Developing a Spanish ASVAB

The ASVAB is a standardized assessment; that is, the ASVAB score scales and testing times are established with regard to a fixed set of test directions, testing conditions, and scoring rules. "Standardization helps to ensure that all test-takers have the same opportunity to demonstrate their competencies." Further, 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 increases with the stakes of the test" (AERA, APA, & NCME, 2014). The standardization is the foundation of the evidence-supported validity and reliability of the ASVAB.

Developing psychometrically equivalent, standardized assessments in multiple languages presents many challenges due to the influence of cultural and linguistic differences that are irrelevant to the intended uses of the tests of interest. Two approaches are often considered in the development of different language versions of an assessment: 1) simultaneous/concurrent development with the source version (e.g., English-language version), or 2) translation and adaptation (transadaptation) of the source version after the source version has been developed.

The International Test Commission's (ITC's) *Guidelines for Translating and Adapting Tests* (International Test Commission, 2017; hereafter, *The ITC Guidelines*) recommend the simultaneous/concurrent development approach of different language versions from the start to "avoid future problems with translating/adapting the source version. . . . At the very least, [developers should] design the source version that enables future translations and avoids potential problems as much as possible." Concurrent development of the assessments allows test developers to minimize item and test differential item functioning (DIF) (Solano-Flores, Trumbull, & Nelson-Barber, 2002).

A Spanish version of the ASVAB would require a transadaptation of the English ASVAB, since the English ASVAB is already well-established and the basis of decades of selection and classification decisions. Test transadaptation, where a test constructed in a source language and culture is transformed for use in a second (target) language and culture, involves both the translation and adaptation of test items and directions written originally in the source language and culture. It often requires more than a direct word-for-word translation and frequently results in the need to replace items and test instructional materials that are unsuitable for transadaptation with items and test directions written in the target language.

Issues about the psychometric equivalence of the original and translated items have been extensively studied. Research has shown that, more often than not, translated items function differently (i.e., display translation DIF) across language groups due to changes in word/item difficulty and differences in cultural relevance, among other factors (Allalouf, Hambleton, & Sireci, 1999). These unintended changes can introduce irrelevant constructs into the translated test and additional variability in test scores (Cohen, Gafni, & Hanani, 2007; Ellis, 1989; Ercikan, 2002; Ercikan, Gierl, McCreith, Puhan, & Koh, 2004), therefore making it impossible for English and Spanish versions of the ASVAB to be psychometrically equivalent.

Linking or relating scores on the target (Spanish) version to scores on the source (English) version of the ASVAB would present another psychometric challenge, with the added disadvantage that it would not produce interchangeable scores (Sireci, Rios, & Powers, 2016). Linking scores across a Spanish ASVAB and the English ASVAB would also have logistical challenges, as it would require a sufficient number of bilingual respondents who could take both language versions of the test. Even if it were logistically feasible to conduct a linking study, scores would not have the same meaning due to differences in the constructs being measured. Because scores could not be used interchangeably across the two language versions of the ASVAB, different evaluation and classification standards would likely need to be developed and used for each version.

The above matters would need to be considered, and relevant processes would need to be followed every time a new English-language version of the ASVAB was developed. The sheer magnitude of work that would be required to transadapt new CAT-ASVAB item pools and new P&P-ASVAB forms into a Spanish-language counterpart would make it impossible to regularly update ASVAB content in a timely fashion.

Additional Considerations

Producing a legally defensible transadapted Spanish version of the ASVAB would present many technical challenges, and would be complex and time consuming, especially for a high-stakes assessment like the ASVAB. It is also a controversial practice because many factors can or may affect the psychometric quality of a transadapted assessment. There are practices recommended by the *ITC Guidelines* that *may* control or reduce *some* of the effects that can threaten the validity and reliability of the transadapted tests. However, following the *ITC Guidelines* would not ensure that scores could be used and interpreted the same way across English and Spanish versions of the ASVAB. Additionally, legal considerations would need to be made regarding whether the ASVAB should also be provided in alternative languages beyond Spanish.

Before any operational administration of a Spanish ASVAB could take place, validity evidence would be required to determine whether scores from the transadapted test are comparable to the English ASVAB and whether the same evaluation standards could be applied to all test-takers (*Standards*, 3.12). Collecting the necessary evidence, such as test-takers' follow-on training and job performance, would take substantial time and effort. Additional time and effort would be required if the validity evidence suggested that separate standards were needed for the Spanish language version.

Producing a Spanish version of the ASVAB would necessitate developing new normative scores and cut scores. A norming study similar to the one conducted with the 1997 Profile of American Youth nationally representative sample would need to be conducted with an updated nationally representative population that would include Spanish-speaking, English-language learners. New normative scores would then need to be developed and a new score scale would need to be introduced to ensure the continuance of the current ASVAB score scale so military policy planners could 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).

Conclusion

Attempting to develop a Spanish version of the ASVAB through a transadaptation process would involve complex, challenging, and prolonged psychometric tasks. Developing an assessment for only the non-native English, Spanish-speaking population, which the transadapted test would target, would bring complications that could hinder or prevent the development of a legally defensible, fair, and valid assessment. The likelihood of large variations in test-takers' Spanish proficiency and literacy levels and the changes in the population from year to year would threaten the validity and reliability of the transadapted test.

The ITC Guidelines call for substantial sample sizes of data for evaluating test items and establishing score accuracy and comparability during the transadaptation process. Acquiring high-quality samples of data with examinees that are bilingual in Spanish and English would be difficult to achieve. It would likely take years to collect sufficient high-quality data.

Validity evidence, as required by the Standards to demonstrate comparability of scores from the two language versions, requires well-designed research studies that would result in years of data collection. Without such evidence, operational use of a transadapted test will not be possible. Without adequate high-quality data, the likelihood that score comparability would be demonstrated is questionable.

Additionally, statistics show that the Hispanic or Latino population in the nation has changed significantly since 2000, and those younger than 33 are mostly English proficient (Pew Research Center, 2016).

Therefore, OPA concludes that the effort, costs, and legal concerns of developing a Spanish-language ASVAB (with potentially questionable results) significantly outweigh any perceived benefits that the assessment would bring. As a result, OPA does not recommend that the Accession Policy Directorate of the Office of the Under Secretary for Personnel and Readiness change the current policy of administering the ASVAB in English only.

References

Allalouf, A., Hambleton, R. K., & Sireci, S. G. (1999). Identifying the causes of DIF in translated verbal items. *Journal of Educational Measurement, 36*(3), 185–198.

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 (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Campbell, J. P., & Knapp, D. T. (Eds.). (2001). *Exploring the limits in personnel selection and classification*. Lawrence Erlbaum.

Cohen, Y., Gafni, N., & Hanani, P. (September 2007). *Translating and adapting a test, yet another source of variance; The standard error of translation*. Paper presented at the annual meeting of the IAEA, Baku, Azerbaijan.

Department of Defense for Personnel and Readiness (2017). *Population representation in the military services: Fiscal year 2017 summary report*. Retrieved from Center for Naval Analyses website: https://www.cna.org/pop-rep/2017/summary/summary.pdf

Ellis, B. B. (1989). Differential item functioning: Implications for test translation. *Journal of Applied Psychology, 74*, 912–921.

Ercikan, K. (2002). Disentangling sources of differential item functioning in multi-language assessments. *International Journal of Testing, 2*(3), 199–215.

Ercikan, K., Gierl, J. J., McCreith, T., Puhan, G., & Koh, K. (2004). Comparability of bilingual versions of assessments: Sources of incomparability of English and French versions of Canada's national achievement tests. *Applied Measurement in Education, 17*(3), 301–321.

Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. Spielberger's, *Adapting educational and psychological tests for cross-cultural assessment*, 3–38. Lawrence Erlbaum Publishers.

International Test Commission. (2017). *The ITC guidelines for translating and adapting tests, second edition* [online]. https://www.intestcom.org/page/16

Krogstad, J. M., Stepler, R., & Lopez, M. H. (2015). *English proficiency on the rise among Latinos: U.S. born driving language changes*. Retrieved from Pew Research Center website: https://www.pewresearch.org/hispanic/2015/05/12/english-proficiency-on-the-rise-among-latinos/

Lopez, G. & Patten, E. (2015). *Hispanics of Puerto Rican origin in the United States, 2013: Statistical profile.* Retrieved from Pew Research Center website: https://www.pewresearch.org/hispanic/2015/09/15/hispanics-of-puerto-rican-origin-in-the-united-states-2013/

Maier, M. H. & Sims, W. H. (1986). *The ASVAB score scales: 1980 and World War II* (A-9). Center for Naval Analyses.

Pew Research Center (2016). *Rise in English proficiency among U.S. Hispanic is driven by the young* [online]. http://www.pewresearch.org/fact-tank/2016/04/20/rise-in-english-proficiency-among-u-s-hispanics-is-driven-by-the-young/

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

Sireci, S. G. (1997). Problems and issues in linking tests across languages. *Educational Measurement: Issues and Practice, 16*, 12–19.

Sireci, S. G. (2005). Using bilinguals to evaluate the comparability of different language versions of a test. In R. K. Hambleton, P. Merenda, & C. Spielberger's, *Adapting educational and psychological tests for cross-cultural assessment*, 117–138. Lawrence Erlbaum Publishers.

Sireci, S. G., Rios, A. J., & Powers, S. (2016). Comparing scores from tests administered in different languages. In N. J. Dorans & L. L. Cook's, *Fairness in educational assessment and measurement*, 181–202. Routledge.

Solano-Flores, G., Trumbull, E., & Nelson-Barber, S. (2002). Concurrent development of dual language assessments: An alternative to translating tests for linguistic minorities. *International Journal of Testing, 2*(2), 107–129.

Welsh, J. R., Kucinkas, S. K., & Curran, L. T. (1990). *Armed Services Vocational Battery (ASVAB): Integrative review of validity studies* (Technical Report No. 90-22). Air Force Systems Command.