Bryson Jones

Carson Streiff

Kyle Rivera

Gamali Felix

Dr. Sipka

Prof. Beardsley

SLC 494

4/28/2024

International Test Development: Localizing US Experiences

Introduction

With Serbia in the process of piloting the official version of their new state standardized testing for high school students, we have examined the standardized testing performed in the United States to determine how to best assist the implementation of the state standardized test in Serbia. In order to do so, we have examined standardized testing in the United States through both the SAT and ACT test. The paper will reflect our findings and research as well as provide recommendations for how standardized testing in Serbia can be improved upon in the future. The paper will highlight how effective the standardized tests in the US have been and how they are administered. Also how these tests affect multilingual test takers and ethical issues that have resulted from such testing and how they can be avoided.

Recommendations

* Prevent teachers from skewing or falsifying test results.
* Do not use test scores as the only means of assessing college readiness.
* Ensure that racial and socioeconomic disparities do not occur.
* Explain tricky vocabulary
* Test spoken and everyday Serbian
* Offer online or quiet in-person options
* Culturally neutral questions
* Adapt for multilingual test-takers
* Continuously update tests based on feedback
* Restricting or limiting the language part of the test
* Have the test online, over paper
* Follow 2-part reading/writing format from SAT
* Incorporate language-specific glossaries
* Consider document-based questions

Key Judgements

When it comes to standardized tests conducted in the United States, the actions of teachers and other educators have a profound influence on the results of the test. Often, they can bias or skew the results, which makes it impossible to get an accurate gauge of the students being tested. Prior to the testing period, teachers in the US are often engaged in the practice of, “teaching to the test”. This practice is when educators craft their lessons and curriculum to cover only the content that students will be tested on.

This is done to raise test scores and demonstrate that students are improving in their education. As a result, the test scores that come from these students are extremely biased. “Teaching to the test, however, is a type of consumer fraud. Any test is only a sample from a large domain of knowledge and skills; mastering a small part of the domain that happens to be tested creates a biased test score. Students, parents, and the public think that more learning has occurred than really did.” (Haladyna 2006). Because of this, unethical test preparation has occurred, which leaves students unprepared for academic advancement. It also diverts time and resources from students to learn about concepts and topics that are not covered by tests.

Additionally, teachers have been caught altering student test scores after the testing period is concluded or have provided hints or directions on how to raise their scores. This is done in order to raise their schools standing and ranking or because teachers and administrators are evaluated based on how well their students perform based on these tests. Meaning that if students underperform, then teachers could be terminated or necessary funds will be withheld from such schools, unless the scores improve. Such an environment creates stress and pressure on both teachers and administrators to obtain the targeted scores in order to avoid the penalties that would result from this outcome.

Another aspect of applying the scores produced by standardized testing in the United States is that they should not be the only information considered for admission into higher education. Both the SAT and ACT are extremely reliable in producing consistent results among test takers. The tests provide administrators with a gauge of the ability of test takers to understand the concepts they are being tested on. This is why many of the champions of these tests continue to push for their usage. However, with both the SAT and the ACT, they are very limited in both scope and focus. They do not cover all areas of a student’s education.

Not only that, but they are also not an accurate metric for determining how well a student will perform at university. A student may score high on either the SAT or ACT and yet receive poor grades in their course work. The opposite is also true. Having test scores be the primary standard for admission into university leaves administrators with no clear guarantee that they will do well achieve the same results in their coursework. Because of this, it is crucial that college admissions offices examine other areas of a student’s application information.

A student’s grade point average (GPA) along with class rankings provide a much more accurate picture of how well they will perform at university. The student’s GPA demonstrates a student's academic performance and how their discipline in their studies translates to their time at university. While the scores derived from either the SAT or ACT can be used to predict a student’s performance in their first year of college, that ability to continuously predict and estimate their achievements loses its power moving further into their time at university.

The last ethical implication to be concerned with for standardized testing conducted in the United States is that the SAT and ACT have often perpetuated racial and socioeconomic disparities across the country. Recent studies have found that students that come from wealthier families score higher than poorer families. Specifically, “students with a family income of $100,000 or more are more than twice as likely as students with family income under $50,000 to have combined scores of 1400-1600.” (Kantrowitz 2022). Beyond this, white students were found to score three times higher than black students and twice as high as Hispanic students. Such wide gaps between different populations of students result in an increase in the likelihood that these students who receive lower scores might be denied entry in higher education based on the outcome of the test.

Additionally, the disparity between genders was also reflected by the scores of both men and women that took the tests. Studies have found that, “matching almost 47,000 men and women on the type of math course taken and grade received, the authors found that, on average, women had scored about 33 SAT points lower than men who had taken the same course and received the same grade.” (Wainer & Steinberg 1992). With such results being the basis for admittance into college, such a gap between men and women’s score would result in less women being able to access secondary education.

The fairness of standardized tests like the US SAT has come under scrutiny. Concerns exist that these tests favor certain student demographics. Do they truly measure ability, or simply reward pre-existing advantages? Multilingual students face another hurdle - the test itself isn't their native language. How can we create a level playing field for all test-takers?

Building a better test requires considering language inclusivity. The Serbian equivalent of the SAT, for example, should encompass all regional dialects and speaking styles. This ensures the test reflects the true spectrum of the language. Standardized tests often focus heavily on reading and writing. Fluency, however, also involves listening and speaking skills. Including these aspects creates a more well-rounded assessment.

Accessibility is paramount for multilingual test-takers. Offering adaptations or translations can level the playing field. This allows them to showcase their actual understanding of the concepts being tested. Transparency and guidance are also crucial. Providing practice tests and helpful tips allows students to prepare effectively. This reduces test anxiety and ensures everyone has a fair shot at demonstrating their abilities.

With specifics related to the implementation of the test, we have a few recommendations related to the SAT that we would propose to make the test relatively thorough for the students. First, we recommend the 2-part system that the SAT uses in which one (reading) section focuses on passages with multiple-choice questions relating to comprehension, analysis, reasoning and integration of ideas from texts and graphs, and the second (writing) section with questions that pertain to passages with issues with grammar, clarity or style. A link to a practice test of this format can be found in the works cited section.

Additionally, more recent versions of the SAT that are tailored to ESL students have included language-specific glossaries that help with key translations. Also, language-specific instructions are included and may be critical in ensuring fairness between students of different languages.

Furthermore, a consideration of document-based questions could be relevant, and would provide students with a chance to read through a longer passage of material rather than shorter, multiple choice questions.

Finally, continuous improvement is essential. Regularly gathering feedback from students and educators is crucial. This allows test developers to identify and address any lingering biases or areas for improvement. The goal is to create a test that accurately measures knowledge and skills, not socioeconomic background or native language. By addressing these issues, we can move towards standardized tests that are truly fair and effective for all students.

Limitations

We recognize that the scope and scale of this report is limited. Therefore we highlight the limitations that have influenced the report and the focus of the research conducted.In coordination with colleagues in Serbia, we focused our research based on their recommendations and on the topics that they desired more information on. Our colleagues in Serbia supplied our team with detailed accounts of the current testing situation in the country and how the new devised test would streamline the testing process for the country. In addition they explained the testing topics and gave personal accounts of test taking in Serbia.The correspondence between our team and our Serbian colleagues shaped the report and provided us with much needed contest for the recommendations and research conduct.

Also, because the researchers of this report are students in an ethics in language testing class at Arizona State University, the focus on the report only deals with the implications of language testing and how it functions in the United States. The ethical implications of the standardized testing conducted in the United States influenced how we approached each of the related topics.Having spent an entire semester studying the ethical issues that form as a result of the standardized testing in the United States, we were cognizant of how we can help Serbia avoid the same issues within their own implementation of standardized testing.

As a result of one of the colleagues that we have worked with in Serbia being unable to disclose information on the formation and operation of the country’s new state standardized testing, we have operated on the limited information that they were able to provide. They provide our team a general understanding of the topics and structure of the test as well as how it would be implemented across the nation, once it is ready for full scale implementation. This information shaped our recommendations and how we understood the testing culture and implementation in Serbia.

Strategic Analysis

With respect to the Serbian Ministry of Education’s attempts to develop a new standardized test it is important for the test developers to avoid the ethical issues that can occur through such tests. In order to ensure that the scores that result from the test are valid it is crucial for them to put measures in place to decrease the likelihood of teachers biasing the results by cheating. Having independent proctors, who have no incentive to bias or alter the scores of test takers would ensure that this issue does not occur during the testing period.

Also, with test scores being a crucial focus of admission into university in the United States, it is important for universities in Serbia to recognize that this should not be the only factor taken into account when reviewing a person’s application. With the Serbian Ministry of Education leaving colleges the freedom to require students to take an entrance exam before they can be considered for admittance. This would give individuals greater access to higher education, without having them to worry of being denied admission on the basis of one test score. Additionally, these universities should review the applicants grades and accomplishments to determine whether or not a person should be accepted into the university. Both of these processes could help Serbian universities provide equitable access to education to all Serbian citizens.

Effective language tests for Serbia require acknowledging the unique aspects of the language and its speakers. Serbian boasts a rich tapestry of regional dialects. A fair test should encompass the variations in vocabulary, grammar, and pronunciation across these dialects. This ensures everyone, regardless of their origin within Serbia, has an equal opportunity to demonstrate their proficiency.

Formal book learning isn't the only path to fluency in Serbian. Many people in Serbia learn the language through daily conversations. The test should therefore assess comprehension and spoken communication skills used in everyday life, not just focus on textbook Serbian.

Test environments can significantly impact performance. Offering the option of online or in-person testing caters to individual preferences. Additionally, ensuring a quiet and focused environment, regardless of the chosen format, helps alleviate test anxiety and allows everyone to perform to their best ability.

By incorporating these considerations, language tests in Serbia can become a more accurate reflection of an individual's true grasp of the language. This promotes fairness and inclusivity, ensuring everyone has a fair shot at demonstrating their Serbian language skills.

Standardized testing is important for assessing students' academic achievements and potential, but it is critical to continuously innovate and enhance these exams to maintain fairness and accuracy. This study proposes two proposals for improving standardized tests: restricting or limiting the language portion of the test and moving to online testing platforms.

For students, particularly those who do not speak the test language natively. In areas with various linguistic backgrounds, this can be a substantial impediment to equal assessment. We can address this issue and make the assessment more accessible to all students by restricting or limiting the language portion of the test. One strategy is to reduce reliance on complicated language passages and questions that demand substantial reading comprehension. Instead, concentrate on evaluating core concepts and talents that transcend linguistic limitations. This could include employing more visual aids, diagrams, and numerical puzzles that are easy to understand regardless of language proficiency. Reducing the language component allows us to devote more time and resources to other portions of the test, such as math, science, and critical thinking. This allows for a more comprehensive evaluation of students' abilities while also leveling the playing field for non-native speakers.

Moving standardized tests to online platforms has various advantages over paper-based assessments. Online testing provides better flexibility, accessibility, and efficiency when administering and scoring exams. It also allows for customization and adaption to fit the demands of different student demographics.

One of the most major advantages of online testing is the ability to deliver examinations in several languages. Allowing students to choose their preferred language allows us to accommodate linguistic variety while also reducing the language barrier that affects performance on traditional paper assessments. This fosters diversity and guarantees that all students have an equal opportunity to demonstrate their knowledge and abilities.

Furthermore, online testing streamlines the assessment process and shortens the time required for grading and score reporting. Students obtain their findings faster thanks to automated scoring algorithms and prompt feedback mechanisms, allowing for timely interventions and support when needed.

Transitioning to online testing necessitates addressing technological issues and ensuring fair access to technology and internet connectivity. With adequate preparation and infrastructure investment, online testing has the potential to transform standardized tests and improve the overall testing experience for both students and educators.

Finally, by incorporating these novel concepts, we can improve the efficacy and fairness of standardized testing. Restricting the language component and moving to online platforms solve crucial challenges like language barriers and accessibility while increasing assessment efficiency and accuracy. These programs are consistent with the purpose of giving all students equal chances for academic success and progress.

Conclusions

In conclusion, to assist the Serbian Ministry of Education as it has begun to restructure its existing standardized test structure, we examined how such tests are conducted in the United States and found these recommendations to be implemented after the trial period for the new state matura is complete. Our recommendations centered around the four areas of the effectiveness of the U.S. SAT and ACT tests, the administration of the SAT and ACT, the experiences of multilingual test takers and the ethical issues that have resulted from such tests. By adapting to the test to account for the recommendations to boost the efficiency of the state matura as well as to ensure that multilingual students are not adversely affected by the test, the Ministry of Education will have necessary information based on experiences from the United States to draw upon in order to improve the test later on. Additionally, that can escape many of the ethical issues that have resulted from these tests in the U.S. as well.

Works Cited

Berry, C. M., & Sackett, P. R. (2009). Individual Differences in Course Choice Result in Underestimation of the Validity of College Admissions Systems. Psychological Science, 20(7), 822-830.<https://doi.org/10.1111/j.1467-9280.2009.02368.x>

Chen, G. (2023, May 10). *When teachers cheat: The standardized test controversies*. Public School Review. <https://www.publicschoolreview.com/blog/when-teachers-cheat-the-standardized-test-controversies>

Djigunović, J. M. (2019, August 7). *affect and assessment in teaching L2 to young learners*. De Gruyter. <https://www.degruyter.com/document/doi/10.21832/9781788924825-005/html>

Haladyna, T. M. (2006). Perils of standardized achievement testing. <https://files.eric.ed.gov/fulltext/EJ750641.pdf>

Kantrowitz, M. (2022, November 9). *How admissions tests discriminate against low-income and minority student admissions at Selective Colleges*. Forbes. <https://www.forbes.com/sites/markkantrowitz/2021/05/21/how-admissions-tests-discriminate-against-low-income-and-minority-student-admissions-at-selective-colleges/?sh=3747afdf3cc1>

Mattern, K. D., & Patterson, B. F. (2013). Test of slope and intercept bias in college admissions: A response to Aguinis, Culpepper, and Pierce (2010). *Journal of Applied Psychology*, *98*(1), 134-147.<https://doi.org/10.1037/a0030610>

Sackett, P. R., Kuncel, N. R., Beatty, A. S., Rigdon, J. L., Shen, W., & Kiger, T. B. (2012). The role of socioeconomic status in SAT-grade relationships and in college admissions decisions. *Psychological Science*, *23*(9), 1000-1007.<https://doi.org/10.1177/0956797612438732>

Sat Practice Test 1–Digital. (n.d.). <https://satsuite.collegeboard.org/media/pdf/sat-practice-test-1-digital.pdf>

*Supports for English learners: Spring 2024 testing - SAT suite of assessments*. Supports for English Learners | Spring 2024 Testing - SAT Suite of Assessments. (n.d.). <https://support.satsuite.collegeboard.org/spring-2024-testing/accommodations/supports-english-learners>

Wainer, H., & Steinberg, L. S. (1992). Sex differences in performance on the mathematics section of the Scholastic Aptitude Test: A bidirectional validity study. *Harvard Educational Review*, *62*(3), 323-336.<https://doi.org/10.17763/haer.62.3.1p1555011301r133>

Zwick, R., & Sklar, J. C. (2005). Predicting College Grades and Degree Completion Using High School Grades and SAT Scores: The Role of Student Ethnicity and First Language. American Educational Research Journal, 42(3), 439-464. <https://doi.org/10.3102/00028312042003439>

“How to Pace Yourself on the SAT.” *The Princeton Review*, www.princetonreview.com/college-advice/how-long-is-the-sat. Accessed 28 Apr. 2024.

“<https://www.princetonreview.com/college-advice/how-long-is-the-sat#:~:text=The%20SAT%20clocks%20in%20at,%2C%205%20minutes%20with%20breaks>

board, college. “How Long Does the SAT Take?” *College Board Blog*, blog.collegeboard.org/how-long-does-the-sat-take#:~:text=The%20total%20time%20for%20the,minute%2035%20seconds%20per%20question. Accessed 28 Apr. 2024.

<https://blog.collegeboard.org/how-long-does-the-sat-take#:~:text=The%20total%20time%20for%20the,minute%2035%20seconds%20per%20question>