

## **Measuring Success in Education The Role of Effort on the Test Itself** (Gneezy et al., 2019)

Standardized tests, like the SAT and the Programme for International Student Assessment (PISA), are used to measure high school students' aptitudes in math, reading comprehension and more. Governments use standardized tests to measure the level of education of the country. Do standardized tests measure skills, effort or external motivations of students?

### **One Sentence Summary**

American students have better test scores when facing external motivations such as monetary reward after the test. Meanwhile, Chinese students have the same performance whether they receive a reward or not. Therefore, there are different motivational factors between cultures making low-stakes standardized tests a poor measurement of a country educational level.

### **Main Findings**

In the PISA<sup>1</sup> math test of 2012 the U.S. high school students ranked number 36 out of 65 countries. Chinese students were number one. How could two countries with similar levels of GDP<sup>2</sup> have a tremendous difference in the PISA rankings? The authors think that the motivation of students to perform well can explain this puzzle.

The SAT defines the future educational career of students in the U.S. Therefore, the students will put an extra effort to perform well. However, a student will not be motivated to perform well in an international standardized test like the PISA because it does not directly affect them. Performing badly in the SAT can greatly reduce future earnings. Meanwhile, a bad performance in the PISA tests will have no discernable short-term effect. Thus, the SAT could be classified as a high-stakes test and the PISA tests as a low-stakes for U.S. students.

The authors measured Chinese and U.S. student's performance on a math test based on PISA test of 2012. Note that they chose a low-stakes tests to measure how motivation influences performance. Before the test, some students (treatment group) are given 25\$, or equivalent in Chinese currency. For each incorrect answer the experimental administrator will take 1\$ from each student<sup>3</sup>. At the same time, some students, in both countries, were not given the monetary incentive (control group). The U.S. students who received the monetary incentive performed better than those who did not. However, the monetary incentive did not have an effect on Chinese students. Chinese students have the sufficient motivation to perform well in both high and low stakes tests. Contrarily, U.S. students need external motivation to perform well in low-stakes tests. In a sense, they only value educational excellence when their performance has an impact on their individual life.

---

<sup>1</sup> For more information about the PISA tests refer to the following [link](#).

<sup>2</sup> On 2019 the United States ranked number one in GDP per capita. China followed them in number 2.

<sup>3</sup> This incentive scheme was carefully designed to maximize the subject's effort in the test.

## **Concluding Remarks**

Governments use public policies to solve complex social problems. It chooses one or several statistics to measure the success of a policy. However, most of the time this statistics fail to measure what the policy aims to fix. In this case, low-stakes standardized tests may not work when comparing educational level across countries with different cultures. However, a complete redesign of the PISA test is not needed. If the PISA tests uses a similar motivational structure it could correctly measure skills instead of motivation.

## **References**

- Gneezy, U., List, J.A., Livingston, J.A., Qin, X., Sadoff, S., Xu, Y., 2019. Measuring Success in Education: The Role of Effort on the Test Itself. Am. Econ. Rev. Insights. <https://doi.org/10.1257/aeri.20180633>.