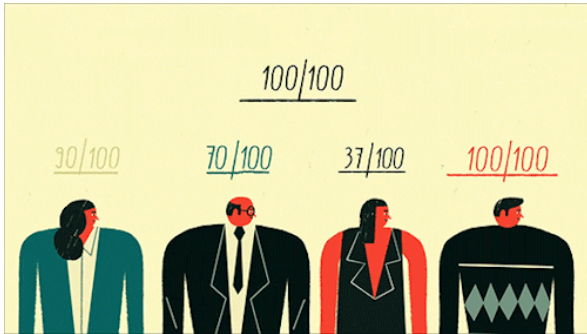


Should we get rid of standardized testing?

Standardized tests are used around the world for everything from evaluating stair climbs for firefighters in France to language examinations for diplomats in Canada to students in schools.



Some standardized tests measure scores only in relation to the results of other test takers. Others measure performances on how well test takers meet predetermined criteria. So the stair climb for the firefighter could be measured by comparing the time of the climb to that of all other firefighters. Similarly, the diplomat might be measured against other test-taking diplomats, or against a set of fixed criteria, which demonstrate different levels of language proficiency.



Although standardized tests are sometimes controversial, they're simply a tool. As a thought experiment, think of a standardized test as a ruler. A ruler's usefulness depends on two things. First, the job we ask it to do. Our ruler can't measure the temperature outside or how loud someone is singing. Second, the ruler's usefulness depends on its design.



Say you need to measure the circumference of an orange. Our ruler measures length, which is the right quantity, but it hasn't been designed with the flexibility required for the task at hand. So, if standardized tests are given the wrong job, or aren't designed properly, they may end up measuring the wrong things.



In the case of schools, students with test anxiety may have trouble performing their best on a standardized test, not because they don't know the answers, but because they're feeling too nervous to share what they've learned. Students with reading challenges may struggle with the wording of a math problem, so their test results may better reflect their literacy rather than numeracy skills.



And students who were confused by examples on tests that contain unfamiliar cultural references may do poorly, telling us more about the test taker's cultural familiarity than their academic learning. In these cases, the tests may need to be designed differently.



Standardized tests can also have a hard time measuring abstract characteristics or skills, such as creativity, critical thinking, and collaboration. If we design a test poorly, or ask it to do the wrong job, or a job it's not very good at, the results may not be reliable or valid.



Reliability and validity are two critical ideas for understanding standardized tests. To understand the difference between them, we can use the metaphor of two broken thermometers. An unreliable thermometer gives you a different reading each time you take your temperature, and the reliable but invalid

thermometer is consistently ten degrees too hot. Validity also depends on accurate interpretations of results. If people say results of a test mean something they don't, that test may have a validity problem.



Just as we wouldn't expect a ruler to tell us how much an elephant weighs, or what it had for breakfast, we can't expect standardized tests alone to reliably tell us how smart someone is, how diplomats will handle a tough situation, or how brave a firefighter might turn out to be. So standardized tests may help us learn a little about a lot of people in a short time, but they usually can't tell us a lot about a single person. Many social scientists worry about test scores resulting in sweeping and often negative changes for test takers, sometimes with long-term life consequences.

We can't blame the tests, though. It's up to us to use the right tests for the right jobs, and to interpret results appropriately.

From the TED-Ed Lesson [Should we get rid of standardized testing? - Arlo Kempf](#)