UNIVERSITY OF PUGET SOUND

Standardized Tests

What factors affects how a student does on these exams?

Oliver B Rozon

12/20/2013

Standardized testing is a cornerstone for the educational system. They are used to measure student achievement in school. There are certain factors that are seen to have an effect on how well a student does on their exams. The results show that the percentage of students who are Asian has the largest positive effect on tests scores, while the student teacher ratio and the percentage of African American students are the largest negative influence on test scores.

Introduction

Standardized testing has been a cornerstone of the educational system for a long time. Standardized testing measures a student's progress throughout their education journey. A standardized test can be known to be, as Henry (2007) suggests "an examination made up of uniform items which can be replicated across an entire domain of students, typically by asking short multiple choice questions which can be easily and cheaply scored by machine". The use of standardized testing have expanded after the passage of the 2002 No Child Left Behind Act, causing some observers to state that standardized testing is causing a "teaching to the test" mentality, which undermines America's ability to produce innovators, imagination, and critical thinkers. Others state that these standardized exams are a fair and objective way to measure a student's abilities.

Besides the debate on whether standardized testing is a good indicator of student progress, there is also debate on whether certain aspects in life; physically, mentally, or situational, have an affect on a student's test score. School districts have tried to manipulate and change certain aspects of the educational system to accommodate for these influences. Their reasoning behind these decisions is to change the negative relationship these factors have on test scores.

This paper explores the issue on whether certain factors affect whether a student does well on a standardized test by looking at data from a large subset of Washington State school districts over a nine year period. The tests that will be looked at are the Washington Assessment of Student Learning (WASL) and the High School Proficiency Exam (HSPE). A

regression will be run using variables that were chosen to represent factors that are most commonlylooked at as having the biggest impact on a student's test score.

Literature Review

Academic standards within the educational system are becoming a staple around the world, especially in mathematics and science. In the United States, the undertaking for national standards and assessments began in 1989 between President George Bush and the nations' governors to set national education goals (Ravitch 1995). Yet, the idea of having standards in the educational system has been around even before President Bush enacted these educational goals Since the 1960s, there have been standardized tests in mathematics and science all around the world. Nowadays, standardized testing and student achievement are usually connected, where a higher test score means the student is learning the information and is better prepared for the future.

There has been debate on whether standardized testing has an effect on student learning, if at all. Some like Orlich (2003), Hicks (2005) and Henry (2007) state that standardized testing does not help students; in fact it may have a negative effect in the future. Due to the increase of standards created by the No Child Left Behind Act, the phrase "teaching to the test" comes about, where teachers teach according to the test, not according to what is necessary for future growth. However, Ravitch (1995) asserts that these raised standards are not just for the best students, but for all students. She does clarify that these assessments only solve part of the problems in the American School District, and that more is needed for it to succeed.

Looking at whether standardized exams work or not is not the only issue that parents, educators, researchers, and governments argue about. There are also disagreements about which influences can affect how a student performs on the exams. These influences can range from cultural, physical, or situational factors.

Throughout the educational system, there is talk about having programs to help students of minority do better in school. Wilson (1997) and others have found strong correlation between exam performance and ethnicity. Is there a reason behind this? One plausible explanation is that standardized exams have a cultural bias against students of minority descent. They bear the cost of past discrimination and past inequities in educational opportunities (Bremmer 2007). However, a physiologist professor named Claude Steele from the University of Michigan conducted an experiment between people who are African American and people who are white. Each group had the exact same credentials, each was given the same exam, and the test giver states that this is a measure of their intelligence. The end result was that the white group performed better than the African American group. For the second phase of the exam, the test giver told the test taker that the exam is not a test of their intelligence but instead it is an instrument to study problem solving. The results show that Blacks did just as well as whites (Horne 2009). The exam was exactly the same, yet African Americans performed better in the second phase than the first, with just a change in description.

Even though getting an education can be free, doesn't mean every student will succeed or that every school is created equal. Poverty is still a prime suspect on why some students

attain better test scores than other students. Lower income families have fewer educational opportunities. They live in areas with a smaller tax base, and as a result their schools are of lower quality (Bremmer 2007). For students who are in families that are not in poverty, they attend schools that tend to have better teachers and more resource-rich educational environments. As a result, students who attend these higher quality schools have better test scores than students who attend the lower quality schools. One way some schools measure how many students are in families in poverty is by looking at students that are under the free and/or reduced lunch program (Orlich & Gifford 2006). To qualify, the income of the student's family has to meet the guidelines set forth by the USDA's Food and Nutrition Service agency. If the family income is at or below 130 percent of the poverty level, the students are eligible for free meals, while those families with incomes between 130 and 185 percent of the poverty levels are eligible for reduced-price meals (USDA, 2013).

Class size has always been an issue with teachers. Back in 2011, teachers from the Tacoma Public School District in Washington State did a strike to protest against proposed features in the school district's new contract proposal, one of which was a two student increase in class size. After a week of striking, a settlement was met, with the previous class size limits intact among other things (Yuan & Matsukawa 2011) (Neroulias 2011). Does class size really have an impact on a student's education? Gene Glass and Mary Smith (1978) have taken a look at that question in hopes of answering the question on whether class size matters. From their research, they demonstrate that by reducing the class size, there is an increase in test scores. There is expected to be small achievements when the class size is in the 20-30 pupil range and

large advantages when the class size is below 20 pupils (Glass & Smith 1978). They do however take a step back in that it is up to the reader to decide whether this advantage is worth the cost.

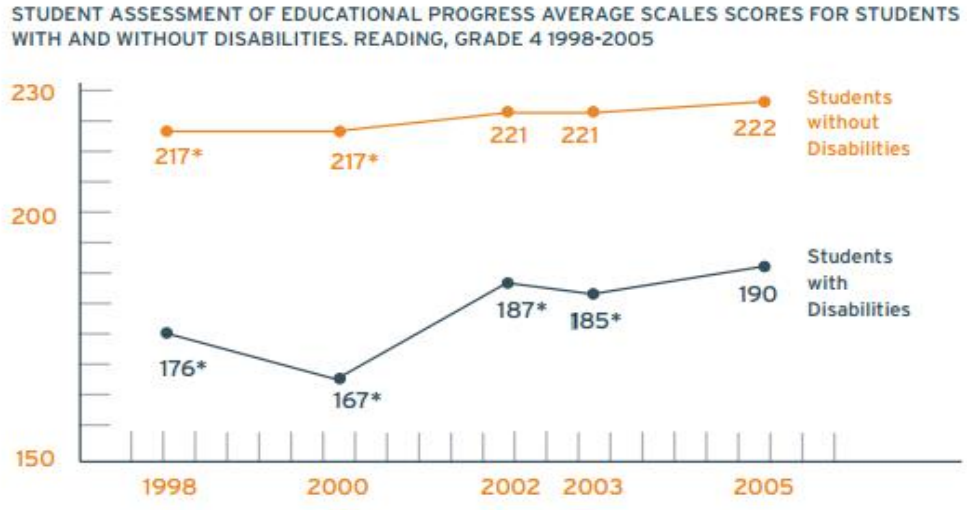
. Teacher experience is thought to have an effect on test scores because teachers with more experience know what works and doesn't work. They use this experience to help the students. There has been debate, however, on whether teacher experience has any affect at all. There is the side that says experience has a big positive effect on student achievement (Clotfelter, Ladd & Vigdor2007). On the other side were people who think there is a weak relationship between experience and student achievement. It is possible that the teachers who have more experience may get students who do poorer in school and act up in class, rather than the students that are more behaved and higher achieving (Fetler 2001). This creates the notion that teachers with experience actually decrease a student's test scores, not increase.

There have been disagreements on how gender plays a role in students and their test scores. According to studies in the past, males score higher in science and math tests while females score high on language and reading tests (Pope& Sydnor 2010). A hypothesis on why males are better at math than females is that males have and developed superior spatial skills and this gives them an advantage in math (Niederle & Vesterlund 2010). Another theory is that during the test, every student is doing about the same. The problem arises when the exam starts to get hard, and the frustration begins to grow. This frustration begins to distract the person, making them less focused and as a result their scores go down (Horne 2009).

Students who have physical or learning problems that hamper their ability to learn and excel in school are known as special needs students. While they take separate classes

specifically for special needs kids, when it comes to standardized exams, there have been some disagreements. Even though special needs students do not have the same learning capabilities as other students, they are required to take the same state exams; where if they do not pass, they do not get their diploma. This causes some arguments on whether these students should take these exams at all. The figure below from the National Center for Learning Disabilities given by Cortiella, C. (2007) gives a comparison between the performances of students without a disability with those who do have a disability.

FIGURE 2: GRADE 4 PERFORMANCE OF STUDENTS WITH AND WITHOUT DISABILITIES ON NAEP READING ASSESSMENT



As the Figure shows, while students with disabilities have had small gains over time, their test scores pale in comparison to those students who have no disabilities. In 2008 California required that special needs students are to pass the state's exit exam to qualify for a

diploma after two years of exemption. This exit exam has about half of all special education students who take it fail, worse than those who are English learners (Asimov 2008). If data suggests that many special needs students are not passing these exams, why are they still forced to take the exam? Jack O'Connell, California state Superintendent, states that this exit exam "is in the students' best interest" (Asimov 2008). The exit exam is designed to test the minimal level of academic skills students need to be successful in the real world. Without it, special needs kids will be at an even greater disadvantage.

Argument

When looking at a standardized exam, there are physical, cultural, and situational influences that affect how a student will score on a standardized test. These factors range from the ethnic backgrounds of the students, poverty, gender, and teacher experience. Furthermore, these factors can have a positive, negative, or no effect on a students test score. This paper aims to show that these factors either have a positive, negative, or no effect on test scores.

Data and Methodology

All the school districts used come from Washington State. Additionally, only the 10th grade exam scores were used in the data set. This is because generally the tenth grade exam scores have an impact on whether the student graduates from high school or not. If the school district did not have a tenth grade score for any of the exams, due to not having a tenth grade or not having enough students take the exams, they were taken out of the data set. The data will span from the 2002-2003 school year to the 2010-2011 school year. The reason behind this time frame is to encompass as many school districts as possible. In the years before 2002, much of the information that is used in this paper was not readily available All demographic and test score data was found in the Office of Superintendent of Public Instruction (OSPI): Washington State Report Card data files.

Regression

TS = f (PASIAN, PAIAN, PBLACK, PHISPAN, PWHITE, PSFRLP, STR, TEXP, PMALES, PSPECEDU, Dummy Variables for each of the school years, with 2002-2003 as base year for dummy variables).

.

The dependent variable is the test score of the school district. In Washington State, tenth grade students are required to pass all four subjects of the state exam: reading, mathematics, writing, and science. . If the student does not pass all four subjects, they are in danger of not graduating.

The first five independent variables all represent the broad idea of race. In this regression, the variables are as followed: Asian, American Indian or Alaskan Native, Black, Hispanic, and White. The data for these five variables is given in percentages. The purpose of these variables is to figure out which race has a bigger impact on how well a student does on his or her test. A negative correlation between minority groups; Black, Hispanic, and American Indian or Alaskan Native, is expected due to cultural, environmental, or structural factors.

The next variable is the percent of students on the free or reduced lunch program. This variable will be used as a proxy for poverty. In short, poverty has a negative effect physically and mentally in terms of test scores. Students who come from higher income families have more access to better resources than students from lower income families. Higher income students can get into better schools, have better means to help them do their school work, and are able to get a decent meal to eat. Mentally, students who are in lower income families hear far fewer words and learn less as babies and toddlers because their parents are out working day and night.

The next variable is the student-teacher ratio. The ratio shows how many students are in the classroom per teacher. The more students there are, the bigger the ratio will be. There is general consensus by teachers that smaller class sizes are what is best for them and for the students. Students receive more individualized attention and interact more with the teacher. On the other hand, people who oppose class reduction do generally acknowledge that there are some benefits as a result of smaller classes, but they argue that the cost of reducing the number of students in the classroom outweighs the benefit from these smaller classes. Even

though the amount may vary in size, generally smaller class sizes do give a positive effect on student test scores .

Experience is one thing that many professions look for when hiring new employees. School districts are no exception. The perceived notion that teachers with more experience are beneficial for students does have credible evidence to support this assumption. However, there is the argument that even though a teacher has more experience, does not mean they know how to teach. Each school year is distinct in their own right, with different kids and different situations. What may have worked one year may not have worked another year. However, in general teacher experience has a positive relationship with how a student does on the exam..

The gender variable will be taken care of by the percent of males in the school district. There is the stereotypical statement that female students don't do as well as male students because of their gender. This statement is supported by the fact that many jobs at this time are male dominated. Others state that gender does not have an effect on how a student does on their exam. This gender variable will be used as a way to examine whether there is a gender bias when it comes to student test taking.

Special Education is a way for school to teach students with special needs in a way that addresses the students' individual differences and needs. They require significant modifications how they are taught and how they learn. They may need extensive remediation, one on one help, a slower-paced curriculum, or other acceptable adjustments. Though special education students learn differently than other students, they may have to take the same standardized

exams as regular students do. This would cause a negative impact on test scores, since the test scores of the special education students would push down the scores of the other students.

Finally, there will be a dummy variable for each school year with the 2002-2003 year as the base year. These dummy variables will take care of any effects that have happened in those years that are hard to quantify. What happened in one year may or may not have happened in another year. Take for example the financial crisis of 2007-2008. While this event happened started during 2007, the aftereffects still linger further on in time. This is something that cannot be felt before 2007.

The regression will have the following general form:

Test Score = $\beta_0 + \beta_1$ Pasian + β_2 Paian + β_3 Pblack + β_4 Phispan + β_5 Pwhite + β_6 Psfrlp + β_7 Str + β_8 Texp + β_9 Pmales + β_{10} Pspecedu + β_{11} D03-04 +β_{12} D04-05 + β_{13} D05-06 + β_{14} D06-07 + β_{15} D07-08 + β_{16} D08-09 + β_{17} D09-10 + β_{18} D10-11

Regression Results Table 1. Coefficients and P-Values of the Independent Variables in relation to the Test Subject

	Reading		Mathematics		Writing		Science	
Independent Variable	Coef.	P-Value	Coef.	P-Value	Coef.	P-Value	Coef	P-Value
Constant	80.193*	0.000	70.320*	0.000	67.761*	0.000	55.930*	0.000
Pasian	0.315*	0.000	0.942*	0.000	0.343*	0.000	0.933*	0.000

Paian	-0.206*	0.001	-0.232*	0.004	-0.111	0.084	-0.121	0.121
Pblack	-0.489*	0.000	-1.025*	0.000	-0.257*	0.011	-0.821*	0.000
Phispan	-0.050	0.378	-0.062	0.423	0.025	0.682	0.013	0.860
Pwhite	0.034	0.551	0.011	0.884	0.078	0.203	0.104	0.162
Psfrlp	-0.201*	0.000	-0.297*	0.000	-0.212*	0.000	-0.299*	0.000
Str	-0.534*	0.000	-0.649*	0.000	-0.343*	0.000	-0.524*	0.000
Texp	0.083	0.401	0.143	0.269	0.183	0.084	0.321*	0.014
Pmales	-0.095	0.271	-0.249	0.026	-0.063	0.490	-0.355*	0.002
Pspecedu	-0.120	0.054	0.029	0.720	-0.134*	0.043	-0.124	0.130
Number of Observations	2056 Used		1829 Used		2061 Used		2005 Used	
R-Squared Value	0.638		0.500		0.691		0.511	

Note: the * means that the coefficient is significant at the 95% level. There was no mathematics test score data for the 2010-2011 school year.

According to these results, the variable that had the biggest positive effect on test scores is the percent of students who are Asian. On the other side, the variable that suggests the biggest negative effect on test scores depends on the exam subject. In the reading and writing portion of the exam, the student-teacher ratio is the biggest influence on a student's test score. For the mathematics and science portion of the exam, the biggest negative influence on test scores is the percentage of black students in the school district. Another observation that can be recognized is that the top two negative factors are the percentage of black students and the student-teacher ratio. The other factors have little to no relationship with test scores.

Even with these results, the data does not suggest that school districts should stop accepting African American and accept more Asian students. It also does not suggest that the school districts should stop accepting students who are in poverty. The underlying suggesting is that the current Washington State educational system is not working in improving the test scores of students in these groups. Granted, improvements generally do not happen automatically. It takes time, money, and a system that will work in the long run. The system may work in one school year, but may fail in the next year. In terms of the time dummy variables, only the 2010-2011 school year in mathematics was insignificant at the 95% level. This does however need to be taken with caution. During the 2010-2011 school year, end-of-course (EOC) exams began to be required by state law. This replaced the math portion of the HSPE. As a result, the OPSI does not have data of the school district's mathematics score for the tenth graders. For three of the four test subjects, the data suggests that the time dummy variables have a generally increasing positive effect on test scores. For the mathematics exam, there starts to be an increasing positive effect on test score, peaking on the 2006-2007 school year, then a decreasing yet positive effect on test scores. In other words, between the years 2003-2007, there is an increasing positive change in their relationship with test scores. After the 2006-2007 school year, there starts to be a constant negative change in the influence on student test scores.

When looking at the R-squared values for each regression, the writing exam regression has the largest R-squared value with a 0.691; this means that about 69.1 percent of the variation of the writing exam scores can be explained by the independent variables. As for the other R-squared values, the smallest is 0.500. This is due to many of the variables to have

insignificant coefficients. This suggests that even though there was more than 1800 observations for each subject exam, the independent variables chosen did not properly explain how they affect test scores.

.

Conclusion

The quest to find the factors that influence a student's test score can be difficult. One researcher's method may be different than other researcher's method, and as a result the results may differ. However, this does not mean that it is impossible. As the data suggests, certain factors like ethnicity, poverty, and the number of students in a classroom have some affect on how well a student scores on an exam. Other factors, such as teacher experience and gender, suggest that there is little to no relationship to test scores. There is room for debate however. The data set, while large in itself, only looks at the school districts in Washington State alone. There may be a change in the results if this same model was used in a different state. In other words, teacher experience may have a significant effect on student test scores in other states.

Bibliography

Asimov, N. (2008, April 2). Disabled Students Required to Pass Exit Exam. *The SFGate.* Retrieved from http://www.sfgate.com/

Bremmer, D. (2007) Determinants of Student Performance on State-Mandated Competency Exams: Evidence from Indiana Middle Schools.

Buddin, R. & Zamarro, G. (2009) Teacher Qualifications and Student Achievement in Urban Elementary Schools. *Journal of Urban Economics,* 66(2), 103-115.

Cortiella, C. (2007) *Rewards & Roadblocks: How Special Education Students are Fairing under No Child Left Behind*. New York: National Center for Learning Disabilities. 17

Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., (2007). How and why do teacher credentials matter for student achievement? NBER working paper 12828.

Fetler, M. (2001). Student Mathematics Achievment Test Scores, Dropout Rates, and Teacher Characteristics. *Teacher Education Quarterly, Winter 2001.* 151-168. http://www.teqjournal.org/backvols/2001/28_1/v28n1_010.pdf

Glass, G., & Smith, M. L. (1978). *Meta-Analysis of Research on the Relationship of Class-Size and Achievement. The Class Size and Instruction Project.* Washington DC: Far West Laboratory for Educational Research and Development.

Henry, P. (2007) The Case against Standardized Testing. *Minnesota English Journal,* 43, 39-71.

Hicks, M. (2005). The Cons of Standardized Testing. http://astro.temple.edu/~mhicks/standardizedtestingcon.html

Horne, E. (2009, Jan 27). The Obama Effect, Perhaps. *Radiolab Podcast.* Podcast retrieved from http://www.radiolab.org/story/91886-the-obama-effect-perhaps/

Margie (2012) Pros and Cons of Standardized Testing. *Bright Hub Education.* http://www.brighthubeducation.com/student-assessment-tools/16137-the-pros-and-cons-of-standardized-testing/

Neroulias, N. (2011, Sept 23) Teachers Ratify Deal Ending Strike in Tacoma, Washington. *Reuters.* Retrieved from http://www.reuters.com/

Niederle, M. & Vesterlund, L. (2010) Explaining the Gender Gap in Math Test Scores: The Role of Competition. *Journal of Economic Perspective,* 24(2), 129-144.

Office of Superintendent of Public Instruction (2013). *Washington State Report card: Data Files.* http://reportcard.ospi.k12.wa.us/DataDownload.aspx

Office of Superintendent of Public Instruction (2013). *How Students in Special Education Participate in State Testing.*

Orlich, D. C., & Gifford, G. (2006). *Test Scores, Poverty and Ethnicity: The New American Dilemma.* http://www.cha.wa.gov/?q=files/Highstakestesting_poverty_ethnicity.pdf

Orlich, D. C. (2003). Longitudinal Effect of the Washington Assessment of Student Learning (WASL) on Student Achievement. *Education Policy Analysis Archives, 11*(18). Retrieved from http://epaa.asu.edu/ojs/article/view/246/372

Pope, D. G., & Sydnor, J. R. (2010) Geographic Variation in the Gender Differences in Test Scores. *Journal of Economic Perspectives,* 24(2), 95-108

Ravitch, D. (1995). *National Standards in American Education: A Citizens Guide.* Washington, DC: The Brookings Institution.

United States Department of Agriculture: Food and Nutrition Service (2013). *National School Lunch Program.* Washington, DC

Wilson, C. A. (1997). Race, Income, and Test Scores: A Structural Model of the Determinants of Test Scores in Toledo Elementary Schools.

Yuan, T. and Matsukawa, L. (2011, August 29). Tacoma Schools, Teacher's Union in Contract Standoff. *King 5 News*. Retrieved from http://www.king5.com/