

## **TEST SCORES, POVERTY AND ETHNICITY: THE NEW AMERICAN DILEMMA**

Donald C. Orlich and Glenn Gifford  
Washington State University  
Pullman, WA 99164-4237

### **Executive Summary**

The authors examined various studies and data sets relating to high-stakes tests and their relationship to family income and ethnicity. Poverty appears to play a major role in depressing test scores with both state sponsored criterion-referenced and national norm-referenced tests. Using zero-order regression, the authors computed a correlation coefficient of poverty to the 2004 SAT total (verbal plus mathematics) and obtained a value of .98 and for 2005 a value of .97. The correlation coefficient of parental income to the 2004 and 2005 ACT composite scores was .99. Correlation coefficients for ethnicity and 2004 data for SAT total were .92 and for the ACT composite .94.

For the SAT, 97% of the variance ( $r^2 = .97$ ;  $p < .001$ ) in test scores may be explained by family income of the test takers. Data for ACT mirror these findings. Given the widespread use of tests to sort and/or classify students, the socioeconomic, social class and ethnicity status of students needs to be analyzed for apparent test bias by the educational community and policy-makers.

### **Introduction**

Discussions about high-stakes tests must address the issue of student poverty. John W. Gardner (1961) observed how schools have long been used to sort and classify children in his classic *Excellence: Can We Be Equal and Excellent Too?* In *Standardized Minds*, Peter Sacks (1999) shows that tests now sort and classify a wide spectrum of groups as well. Discussions

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa “Summit on Public Education,” Washington, D. C.*

about the ACT, SAT and Washington Assessment of Student Learning (WASL) by Achieve, Inc., the Partnership for Learning, the Washington State Superintendent of Public Instruction and Business Roundtables virtually ignore this social phenomenon. Thus, it behooves any advocates of high-stakes tests to at least ask, “What is the impact of student poverty on test scores as a mechanism of sorting and classifying children?”

### **A Short Review of Published Literature**

Ronald C. Nyhan and Mohamad G. Alkadry (1999) statistically analyzed the relationship of class size, expenditure per student and socioeconomic status on student achievement test scores in three south Florida counties. Poverty was the primary determinant of student achievement. A parallel finding was also reported when English, Scottish, Welsh and Irish student test scores were analyzed (McCallum and Demie, 2001).

Mark Hornbeck (2001) wrote that one in five or about 600 Michigan schools would fail to meet that state’s standards. Again, there are correlations with family income. Similarly, Alan H. Schoenfeld (2002) provided data showing that economic status has a negative learning impact on poor children and children of color. Jaekyung Lee’s (2002) analysis of several social factors showed parallel patterns to those above.

**Findings in Denver.** Alan Gottlieb (2002) discussed research commissioned by the Piton Foundation. Results showed high concentrations of low-income children in Denver’s neighborhood schools kept their achievement levels below what they should be. Data illustrated that low-income elementary school children in Denver performed “. . . significantly better on standardized tests when they attend schools where fewer than 50 percent of the students are poor” (p. 1). In Denver, low-income children demonstrate lagging achievement levels. Again,

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa "Summit on Public Education," Washington, D. C.*

the metric for poverty was free and/or reduced lunch (Gottlieb, 2002). The data presented lead to the conclusion that childhood poverty will predict poor achievement scores on high-stakes tests.

In a thought-provoking essay on the impact of poverty on test results and using data from England, Wales and two American states, William A. Firestone and David Mayrowetz (2000) conclude that we would be better served if we determined what constitutes good educational practice rather than emphasizing incentives via high-stakes tests.

**Boston Metro Findings.** Craig Bolon (2001) reviewed mathematics tests scores from academic high schools in metropolitan Boston. His conclusion follows.

“The state is treating scores and ratings as though they were precise educational measures of high significance. A review of tenth-grade mathematics test scores from academic high schools in metropolitan Boston showed that statistically they are not. Community income is strongly correlated with test scores and accounted for more than 80 percent of the variance in average scores for a sample of Boston-area communities: Once community income was included in models, other factors--including percentages of students in disadvantaged populations, percentages receiving special education, percentages eligible for free or reduced price lunch, percentages with limited English proficiency, school sizes, school spending levels, and property values--all failed to associate substantial additional variance. Large uncertainties in residuals of school-averaged scores, after subtracting predictions based on community income, tend to make the scores ineffective for rating performance of schools. Large uncertainties in year-to-year score changes tend to make the score changes ineffective for measuring performance trends” (p.1).

Collectively, these studies suggest that our findings do not represent an isolated phenomenon.

### **The ACT and SAT**

Similar findings tend to be found when examining SAT scores (Fleming and Garcia, 1998; Adelman, 1999-2000; Nairn, 1980). Data published in 2004 regarding the SAT scores of college bound high school seniors form a linear function in which there is a very strongly positive correlation between parental income and students' SAT scores (*Fair Test Examiner*, 2004). Table 1 shows gender, ethnicity and family income data for 2004 ACT college bound seniors. Table 2 shows the same data sets for the SAT.

Using a regression analysis plotting family income vs. SAT total, the proportion of shared variance of was calculated at .97 by the authors. Using the same analysis, family income vs. ACT Composite, the proportion of shared variance was .99 (using zero order regression, bivariate correlation). These correlations show an extremely high relationship. Of course, they are not “causal,” but the data should be of concern to every member of the education community. When using the regression analysis of ethnicity and SAT Composite, the proportion of shared variance was calculated at .92, and for the ACT .93.

The test data from the ACT and SAT illustrate an ethnic component related to achievement on high-stakes tests. These data tend to indicate that scores of various ethnic groups most probably are related also to socio-economic conditions. Poverty and ethnicity appear to be inextricably related. Additionally, David C. Berliner (2005) discusses these points, and his data tend to corroborate the above conclusions.

These data would indicate that the high-stakes test syndrome now in vogue due to the “No Child Left Behind Act” will ultimately cause some serious “backlash” by the minority communities and the poor. Policy-makers tend either to ignore these data, or perhaps are yet unaware of them and their ultimate impact.

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa “ Summit on Public Education,” Washington, D. C.*

**Table 1. College Bound Seniors Average ACT Composite Scores, 2004 and 2005**

	Composite Scores	
	2004	2005
<b>Gender</b>		
Female	20.9	20.9
Male	21.0	21.1
<b>Ethnicity</b>		
American Indian or Alaskan Native	18.8	18.7
Asian-American or Pacific Islander	21.8	22.1
African-American or Black	17.1	17.0
Mexican-American, Chicano, Latino	18.4	18.4
Puerto Rican, Cuban, Other Hispanic	18.8	18.9
Caucasian-American, White	21.8	21.9
Multiracial	20.9	20.9
Other	19.4	19.4
Prefer Not to Respond (3%)	22.0	22.0
No Response (4%)	20.7	20.0
<b>Family Income</b>		
Less than \$18,000/year	18.0	17.9
\$18,000 - \$24,000/year	18.7	18.6
\$24,000 - \$30,000/year	19.4	19.3
\$30,000 - \$36,000/year	19.9	19.8
\$36,000 - \$42,000/year	20.4	20.3
\$42,000 - \$50,000/year	20.9	20.9
\$50,000 - \$60,000/year	21.3	21.3
\$60,000 - \$80,000/year	21.9	21.9
\$80,000 - \$100,000/year	22.5	22.5
More than \$100,000	23.5	23.5
No Response (23%)	21.1	21.1
<b>All Test Takers</b>	<b>20.9</b>	<b>20.9</b>

**Sources:** ACT, *ACT Assessment Results 2004*. Approximately 1.17 million test-takers, of whom 56% were female. Data used with permission of the *Fair Test Examiner*. Retrieved from <http://www.fairtest.org/nattest/ACT%20Scores%202004%20Chart.pdf>.

“Latest SAT, ACT Results Flat.” *Fair Test Examiner*, Fall 2005, 19(4): 4-5. Approximately 1.19 million test takers, of whom 56% were female for 2005.

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa  
 “ Summit on Public Education,” Washington, D. C.*

**Table 2. 2004 College Bound Seniors Test Scores: SAT**

	<b>Verbal</b>	<b>Math</b>	<b>Total</b>
<b>Gender</b>			
Female	504	501	1005
Male	512	537	1049
<b>Ethnicity</b>			
American Indian or Alaskan Native	483	488	971
Asian, Asian American or Pacific Islander	507	577	1084
African American or Black	430	427	857
Mexican or Mexican American	451	458	909
Puerto Rican	457	452	909
Other Hispanic or Latino	461	465	926
White	528	531	1059
Other	494	508	1002
No Response (19%)	522	535	1057
<b>Family Income</b>			
Less than \$10,000/year	422	450	872
\$10,000 - \$20,000/year	440	457	887
\$20,000 - \$30,000/year	459	467	926
\$30,000 - \$40,000/year	478	482	960
\$40,000 - \$50,000/year	493	496	989
\$50,000 - \$60,000/year	501	504	1005
\$60,000 - \$70,000/year	507	510	1017
\$70,000 - \$80,000/year	515	518	1033
\$80,000 - \$100,000/year	527	530	1057
More than \$100,000/year	533	562	1115
No Response (42%)	Scores Not Reported		
<b>All Test-Takers</b>	<b>508</b>	<b>518</b>	<b>1026</b>

**SOURCE:** College Board, *College-Bound Seniors 2004: A Profile of SAT Program Test Takers*. Approximately 1.42 million test takers, of whom 53.8% were female. Data used with permission of the *Fair Test Examiner*. Retrieved from <http://www.fairtest.org/nattest/SAT%20Scoresn%202004%20Chart.pdf>.

**Table 3. 2005 College Bound Seniors Average SAT Test Scores**

	<b>Verbal</b>	<b>Math</b>	<b>Total</b>
<b>Gender</b>			
Female	505	504	1009
Male	513	538	1051
<b>Ethnicity</b>			
American Indian or Alaskan Native	489	493	982
Asian, Asian American or Pacific Islander	511	580	1091
African American or Black	433	431	864
Mexican or Mexican American	453	463	916
Puerto Rican	460	457	917
Other Hispanic or Latino	463	469	932
White	532	536	1068
Other	495	513	1008
No Response (19%)	511	525	1036
<b>Family Income</b>			
Less than \$10,000/year	426	458	884
\$10,000 - \$20,000/year	443	463	906
\$20,000 - \$30,000/year	463	474	937
\$30,000 - \$40,000/year	480	487	967
\$40,000 - \$50,000/year	496	500	996
\$50,000 - \$60,000/year	505	509	1014
\$60,000 - \$70,000/year	511	515	1026
\$70,000 - \$80,000/year	517	522	1039
\$80,000 - \$100,000/year	529	534	1063
More than \$100,000/year	554	565	1119
No Response (42%)	Scores Not Reported		
<b>All Test-Takers</b>	<b>508</b>	<b>520</b>	<b>1028</b>

**SOURCE:** College Board, *College-Bound Seniors 2005: Total Group Profile Report*.

Approximately 1.48 million test takers, of whom 53.0% were female.

Table 3 is from: “Latest SAT, ACT Results Flat.” *Fair Test Examiner*, Fall 2005, 19(4):

### **And in Washington State?**

Martin L. Abbott and Jeff Joireman (2001) published a study examining the relationship of achievement to low income and ethnicity. Their extensive review of literature for the period between 1990 and 2000 demonstrated a definite correlation between poverty and low-test scores. Scores on the *Washington Assessment of Student Learning* (WASL) were analyzed for the total population for whom WASL scores were collected. The proxy for poverty was free and/or reduced lunch. They reported that “reading scores are negatively correlated with the percentage of students on free lunch in a given school ( $r = -.72$ ). . . .” (p. 11). In Table 3 of their study, they reported the following correlations: Math,  $-.68$ ; Listening,  $-.67$ ; and Writing,  $-.60$ . This analysis indicates that as poverty increases scores on the WASL decline.

Simultaneously, and unknown to Abbott and Joireman, in the summer of 2001 the authors (Orlich and Gifford) computed correlations for 26 school districts in the North Central Educational Service District—north central Washington state—also using free and/or reduced lunch as the poverty measure. Using zero-order regression (bivariate correlation) we found the following correlations between poverty and fourth-grade WASL scores: Reading,  $0.51$ ; Math,  $0.55$ ; and Writing,  $0.31$ . These correlations indicate that as poverty increases, failure rate on the WASL increases.

### **A Twist on the Tale of Two Cities**

Let us now introduce selected data showing WASL comparisons of two divergent socio-economic groups. Table 4 illustrates the WASL test scores of one of the State of Washington’s highest family income school districts and all children from low-income homes in the state. Low income is defined in that the children are eligible for free and/or reduced lunch at school. The data are shown simply to illustrate the impact that living in a low-income household has on a



*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa “Summit on Public Education,” Washington, D. C.*

child’s WASL test scores. It is not the intent of the authors to single out Mercer Island for being “upscale.” Nevertheless, the children living on Mercer Island have an abundance of “*Social Capital.*”

Embedded within the social context of the school is a concept called *social capital*. Social capital is broadly defined as a sum of interpersonal relationships that provides support or encouragement. Sources of social capital include families, communities, public institutions, churches, clubs, social relationships and collaborative networks. Children living in poverty have very limited sources of social capital. (See Marzano 2004, and Putnam 2000 for an extended treatment of the concept.)

Table 4 provides one shred of evidence supporting our thesis that socioeconomic factors affect achievement. Examining the data in Table 4, one observes a pronounced and significant difference at every grade level and for every subject. Those advocating the WASL and other high-stakes tests have totally ignored the social context of schooling. This omission must be questioned and those responsible asked, “Why?” (For a detailed examination of the impact of poverty on children, we refer you to Ruby K. Payne (2003), *A Framework for Understanding Poverty*, 3<sup>rd</sup> Edition, Highlands, Texas.)

If one were to calculate the percentages shown in Table 3 as percentiles, then the effect sizes would range from 1.2 to 2.3. These effect sizes (see Cohen 1988, Marzano et al. 2001, p. 160) illustrate extraordinary achievement gaps. The overall conclusion from Table 4 is that the parental income of children in the State of Washington very closely mimics the data already provided from national samples.

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa “ Summit on Public Education,” Washington, D. C.*

**TABLE 4. Percent of Students Passing the WASL By Grade Level: Low Income Children Compared to Students From Mercer Island, 2002-2003.**

<b>Low Income Children</b>			
<b>Grade Level</b>	<b>Mathematics</b>	<b>Reading</b>	<b>Writing</b>
4	40.2	51.8	40.3
7	19.6	29.5	44.7
10	24.1	42.9	44.7
<b>Children From Mercer Island</b>			
<b>Grade Level</b>	<b>Mathematics</b>	<b>Reading</b>	<b>Writing</b>
4	86.8	86.2	85.2
7	80.7	84.0	89.6
10	78.7	83.0	84.1

**SOURCE:** Washington State Report Card. Files of Office of State Superintendent of Public Instruction.

**More About the Path of Poverty.** On March 1, 2005, The United Nations released *Child Poverty in Rich Countries: 2005*, The Innocenti Report Card No. 6. The Nordic countries had the lowest levels of child poverty in the “developed” countries of the world, primarily due to very highly subsidized social benefits paid directly to families. The United States of America and Mexico had the world’s worst child poverty rates. For Mexico, the percentage was 27.7 and for the USA it was 21.9

The report writers stated that such disparity of wealth leaves many children, by no fault of their own, at a social disadvantage. The report also noted, as we have already addressed, that there is a close correlation between poverty and educational underachievement.

Poverty is a powerful force in educational deficits, and you will not find advocates of high-stakes testing discussing that social issue, including Achieve, Inc. or The Business Roundtables. One simply has to ask, “Why the silence?” One could speculate that Americans have been subtly conditioned and misled into believing that childhood poverty is really not our social

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa “Summit on Public Education,” Washington, D. C.*

problem. This condition is an individual problem with slackers who refuse to pull themselves up by their own bootstraps to achieve the American Dream. And let us not forget the anti-poverty messages carried by the nation’s media during the 1970’s – 1990’s about all those “welfare Queens.” In a nation that proclaims to defend democracy against all comers, it politely turns its back on its own poor.

This is an opportune moment to reflect on Michael Harrington’s (1962) *The Other America: Poverty in the United States*. His examination of the nation’s conscience appeared at a time when we were celebrating the achievements of post World War II. Harrington argued and provided substantive data that several millions of Americans were trapped in a culture of poverty.

### **A Theoretical Consideration**

Results of two independent studies have demonstrated the relationship of student poverty to WASL scores. These were congruent with national and international reports on high-stakes test scores. Thus, these are not an isolated phenomenon. There appears to be a pattern showing that a child’s socioeconomic status may be used as a predictor of success or failure on high-stakes tests. Applying the *fairness doctrine*, this is a strong indictment against the WASL as a worthwhile assessment. Policymakers are penalizing children for conditions over which these youngsters have no control.

What are some attributes of childhood poverty negatively impacting schooling outcomes? The first is inadequate nutrition as a definite relationship exists between malnutrition and cognitive dysfunction. Chronic Vitamin B1 and B12 deficiencies have a proven relationship to cognitive impairment (Beers and Berkow, Eds., 1999).

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa “Summit on Public Education,” Washington, D. C.*

Contributing factors may be single-parent family structure and poverty (Shim, Felner and Shim, 2001; *kids count DATA BOOK, 2005*). The resource allocation for schools is directly related to property tax bases, thus persons of poverty do not have the fiscal support to provide optimal learning environments for their children. While we did not collect data on these attributes of poverty, we present them as hypotheses to be tested.

**The WASL: An Aversive Experience?** Jeffrey T. Fouts (2002) released a second and most disturbing finding. He followed the progress of 14,860 4<sup>th</sup> grade students in 1998 with the 12,827 7<sup>th</sup> grade students in 1998 on the WASL. He reported in 2002 “In total, of students who scored at Level 1 reading [flunk] in 1998 (4<sup>th</sup> grade), only 3.1% of those students were able to meet the standard (Level 3 or Level 4) [pass] by 2001 (7<sup>th</sup> grade)” (p. iii). These data call into question the utility of the WASL. Fouts then reported “. . . for every 100 students who score at 4<sup>th</sup> grade Level 1 in reading and proceed through the *current* Educational system, approximately 34% or 34 of those students will meet the reading standard [Level 3 or 4] when they reach 10<sup>th</sup> grade” (p. iii).

The above clearly demonstrates that the WASL serves only as a summative measure and provides little tangible evidence that student achievement is improved. This conclusion is supported by Orlich’s study (2003) showing very little or no yearly effect on student achievement as a consequence of administering the WASL. The feedback from the WASL is a number: 1 and 2 (fail: did not meet standard); 3 and 4 (pass: met standard). Robert Marzano et al. (2001) showed an effect size of  $-0.08$  when feedback is simply pass or fail. We would posit that such feedback is totally inappropriate to aid student learning, just as the negative effect size shows. Comprehensive and timely feedback are two factors that are known to influence achievement in a positive fashion. In 1992 and after synthesizing 8,000 studies, John Hattie wrote that, “The

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa “Summit on Public Education,” Washington, D. C.*

most powerful single modification that enhances achievement is feedback” (p. 9). Feedback factors are missing from the WASL.

These results are interpreted as evidence that the WASL is creating aversive educational conditions. Fourth grade students are apparently exhibiting characteristics of *learned helplessness*. As John Cosgrove (2000) notes “Learned helplessness is a specific consequence of one particular form of psychological stress: being given an impossible task” (pp. 45-46).

Writing about the WASL in the *Seattle Post-Intelligencer* of April 24, 2000, Debera Carlton Harrell stated, “But in many places, student meltdowns are preceding academic liftoff.” She noted that there are “concerns about too much stress on children and teachers, and a dearth of resources . . . .” (p. 1).

Harrell’s story was followed by Wendy Harris who wrote a front page feature in *The Spokesman-Review* of May 11, 2000, titled WASL a “ ‘monster’ of a test.” In her article, she cited Wilson Elementary School fourth-grader Alan Guthrie who drew the WASL as “a huge monster that eats children and gets stronger from their fear” (p. 1). These two local stories lead to our next point. Even though anecdotal, these stories illustrate a stress factor.

**Learned Helplessness?** Robert Sapolsky (1994) cites the research of Donald Hiroto and Martin E. P. Seligman showing how learned helplessness is a function of environmental conditions that are adverse. If 4<sup>th</sup> grade children have aversive experiences with the WASL, that is, fail to achieve the arbitrary standard that is set, then as 7<sup>th</sup> graders they probably perceive passing the 7<sup>th</sup> grade WASL as another impossible task.

**The WASL and Maslow’s Hierarchy of Needs.** Let us now invoke the theoretical construct of Abraham Maslow’s (1954) “Hierarchy of Needs.” Maslow constructed eight ascending psychological needs. Level 1 relates to basic physiological necessities such as food and water.

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa “Summit on Public Education,” Washington, D. C.*

Level 2 relates to safety needs, including shelter and security; while Level 3 describes the *belongingness* needs. Level 4 is *Esteem*—the need to achieve, be competent, gain approval and recognition. Children in grade 4 who are informed that their WASL scores are *not meeting standard* are being denied an element of *Esteem*. We argue that being involuntarily subjected to the WASL by the state government of Washington induces the condition of *Learned Helplessness*.

Maslow’s fifth Level is *Cognitive*—the need to know, understand and explore. It is obvious that thousands of 4<sup>th</sup> grade children in Washington State have not entered that level—again based on the WASL test scores. With the *cognitive* need being contingent on the four basic physiological and psychological needs, then it becomes apparent that poverty in general and learned helplessness in specific may be playing a more aversive role in student achievement than previously anticipated.

Children who live in impoverished environments with little chance to feel secure may be unable to ascend beyond Levels 1 and 2 in Maslow’s Hierarchy. This is because they must attend to and focus on their basic living necessities. Additionally, youth in these circumstances tend to be exposed to stress and violence, thus remaining at Level 2 since their safety needs are unmet. Progression to Level 5 is contingent on satisfying the first four levels. Therefore, since these children must attend to the lower levels of Maslow’s Hierarchy, it becomes evident that they cannot attend to matters such as mastering the WASL.

### **Stress, Depression and the “Final” Solution**

It can be predicted that *stress* will develop as a consequence of a combination of learned helplessness and lack of having met the first four needs in the Maslow Hierarchy. *Depression* may set in when a stressful situation is perceived to be hopeless. This being the case, then no

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa "Summit on Public Education," Washington, D. C.*

amount of testing will help children perform any better; probably they will perform worse. We do present two case studies to support the above assertion.

The first piece of evidence comes not from any child, but from the suicide of Betty Robinson, Principal of the Simonton Elementary School in Gwinnett County, Georgia (near Atlanta). Robinson had been the founding principal of Simonton in 1993, but on April 11, 2002, she locked herself inside her school office and shot herself in the head. This action, it was speculated, was a direct result of poor test scores within the school. Her 1,600 pupil school is a Title I school, meaning it has many poor children. Further, the children in her school had not met the arbitrary improvement expectations set by the state of Georgia on its high-stakes test, which is the only measure used in Georgia to assess school success (Hartstein and Jones, 2002).

The authors note the particular circumstances surrounding this suicide. The chosen place and manner of death suggests a cognitive mindset reflecting helplessness and depression over the students' poor test scores. It is reasonable to postulate that this administrator felt that there was no way out and therefore ended her life.

A similar incident took place in Eynesbury, England (CNN, 2000), when a teacher, Pamela Relf, drowned herself leaving a note saying she was upset by her school's poor test performance. Thus, two case studies are presented to illustrate the gravity of high-stakes tests on educators. It is critical to recognize that adults committed these tragic suicides. Adults have a vast array of coping strategies unattained by adolescents or children. One may speculate an increased frequency of such self-destructive behavior will occur within youngsters. Fatal destructive patterns of behavior may occur when children are informed they "flunked" a mandatory test. They lack many psychological resources necessary to cope with stress and depression. This is simply a function of their age and lack of experience.

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa "Summit on Public Education," Washington, D. C.*

**A Modest Prediction.** We predict that if the current trend of using high-stakes tests continues then the schools will become victims of student destructive behaviors. The acts, we speculate, will be in the form of wanton destruction of classrooms and buildings. The social consequences of labeling a generation of adolescents as being flat-out failures from one test needs serious psychological and psychiatric evaluation. Pay heed now or pay later!

And allowing high-school students to take the WASL five times is probably one of the worst educational policies to ever come from Washington's school reform agenda. Is anybody thinking about consequences?

### **Social Class and Schooling Considerations**

While the focus of this paper is on poverty and school achievement, it is necessary to at least consider the general aspects of "social class and schooling." Michael S. Knapp and Sara Woolverton (2004) discuss the topic in detail. In this paper we will simply highlight their key points. Knapp and Woolverton illustrate with data sets and extensively reviews the following findings relative to social class and its effect on education.

1. Social class, prestige and socio-economic ranks impact schooling.
2. Social class is related to the concept of "meritocracy."
3. There is a universal correlation between social class and educational outcomes.
4. The correlations between social class and educational attainment tend to hold over time and across cultures, worldwide.
5. Social class is fluid, not fixed, with education being a strong determinant.
6. Social class and ethnicity tend to be explicit bases for tracking.
7. The economic and social aspects of class affect a student's ability to learn.
8. Poverty plays a detrimental role in student achievement.



*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa “Summit on Public Education,” Washington, D. C.*

Pauline Lipman (2004) analyzed Chicago’s public schools reform, which she described as a business, top-down model. She concluded that the imposed educational and economic policies exacerbated race and socioeconomic disparities. Lipman noted that accountability via high-stakes tests adversely affected teacher morale.

Taken collectively, social class, poverty and ethnicity are factors that must be considered by the U.S. Congress and every state legislature as high-stakes tests become mandated as a sorting mechanism that serves as the primary determinant of high school graduation. The children of the working classes are at risk.

This is an appropriate time to examine Table 5, which illustrates the disparity between various ethnic and social groups on various state sponsored high-stakes tests. These data, again, verify our conclusion that the tests show a bias against selected social groups. Also observe that the state of Washington has the worst “pass” rate of the sampled states. This again is a strong indicator that the WASL has a built-in bias and is developmentally inappropriate.

**TABLE 5. Percentage of High School Students Passing an Exit Exam on the First Try for All Students and By Subgroups**

Student Subgroups	AL		AK		AZ		GA		IN	
	Math 2003	AL Reading	Math 2004	AK Reading	Math 2003	AZ Reading	Math 2003	GA ELA	Math 2003	IN ELA
All	79%	88%	67%	70%	36%	59%	91%	95%	67%	69%
White	86%	93%	76%	82%	49%	74%	94%	97%	73%	75%
Black	66%	79%	44%	58%	21%	44%	78%	89%	33%	39%
Hispanic	73%	72%	53%	61%	18%	38%	79%	79%	46%	44%
Asian	91%	87%	68%	64%	61%	72%	94%	90%	85%	78%
Native American	82%	90%	58%	68%	14%	35%	86%	94%	54%	57%
ELL	74%	68%	38%	29%	7%	13%	68%	61%	41%	27%
Free or Reduced Lunch	68%	80%	47%	46%	NA	NA	NA	NA	46%	48%
Students with Disabilities	46%	59%	23%	22%	10%	26%	52%	69%	27%	22%
Student Subgroups	LA		MD		MA		MN		NV	
	Math 2003	LA Reading	Algebra 2003	MD English	Math 2003	MA ELA	Math 2004	MN Reading	Math 2004	NV Reading
All	68%	71%	53%	40%	80%	89%	71%	81%	43%	77%
White	85%	87%	68%	52%	86%	94%	78%	87%	54%	86%
Black	51%	55%	28%	20%	57%	76%	31%	50%	22%	62%
Hispanic	68%	69%	39%	29%	54%	66%	38%	62%	25%	62%
Asian	87%	77%	76%	59%	88%	88%	58%	63%	53%	81%
Native American	77%	80%	46%	26%	65%	80%	43%	56%	29%	72%
ELL	60%	47%	NA	NA	57%	42%	29%	36%	13%	34%
Free or Reduced Lunch	54%	57%	NA	NA	NA	NA	47%	61%	25%	60%
Student with Disabilities	23%	18%	NA	NA	53%	70%	28%	40%	6%	30%

**TABLE 5. (Continued)**

<b>Student Subgroups</b>	<b>NJ Math 2003</b>	<b>NJ Lang. Arts</b>	<b>NM Math 2003</b>	<b>NM Reading</b>	<b>NY Math 2003</b>	<b>NY English</b>	<b>NC Reading &amp; Math 2002</b>	<b>OH Math 2004</b>	<b>OH Reading</b>
All	66%	80%	81%	89%	83%	85%	78%	68%	79%
White	77%	88%	91%	96%	91%	91%	87%	74%	83%
Black	33%	61%	71%	86%	65%	72%	66%	38%	58%
Hispanic	42%	63%	76%	87%	64%	69%	52%	50%	63%
Asian	83%	87%	94%	93%	89%	86%	77%	84%	84%
Native American	57%	74%	72%	81%	80%	79%	66%	71%	76%
ELL	22%	18%	64%	75%	61%	55%	38%	NA	NA
Free or Reduced Lunch	36%	57%	72%	83%	72%	76%	NA	NA	NA
Students with Disabilities	22%	35%	43%	60%	65%	61%	45%	NA	NA
<b>Student Subgroups</b>	<b>TN Math 2003</b>	<b>TN Lang.</b>	<b>TX Math 2004</b>	<b>TX ELA</b>	<b>VA Math 2003</b>	<b>VA English</b>	<b>WA Math 2003</b>	<b>WA ELA</b>	
All	75%	87%	85%	87%	80%	92%	39%	60%	
White	85%	90%	91%	92%	85%	95%	44%	65%	
Black	52%	78%	73%	82%	65%	86%	14%	37%	
Hispanic	71%	83%	78%	81%	73%	88%	16%	35%	
Asian	87%	90%	95%	91%	89%	94%	47%	64%	
Native American	76%	83%	88%	89%	77%	92%	22%	43%	
ELL	60%	55%	59%	42%	74%	79%	8%	12%	
Free or Reduced Lunch	61%	77%	79%	82%	69%	86%	24%	43%	
Student with Disabilities	41%	43%	55%	56%	51%	70%	4%	12%	

Na = Not Available

Notes: Alaska’s results are preliminary as of May 2004 with district verification still pending. Nevada’s figures for students with disabilities are only for students with Individualized Education Programs under the Individuals with Disabilities Education Act and do not include students with disabilities who are served under Section 504 of the Rehabilitation Act. Texas and Washington, instead of using free or reduced lunch data to determine low-income students, disaggregate data by students eligibility for Title 1, Part A.

**SOURCE:** Center on Education Policy, based on information collected from state departments of education, July 2004. **Source of Table 5:** Center on Education Policy. *State High School Exit Exams: A Maturing Reform*, Table 3, p. 38, 2004.

### **Reasonable Cognitive Expectations**

Our final set of evidence comes from an array of age, grade and developmental levels. Herman T. Epstein (2002) compiled these national and international data on the cognitive development of children at varying ages. The data in Table 6 support our argument that many of the tasks on the WASL and other nationally used tests are beyond the cognitive abilities of most fourth-graders. It is noted that at the fourth grade level a mere one-percent of children was found to be at the entry level for formal operational thinking. Linda Shaw (1999) noted WASL developmental difficulties in a feature story about the WASL. Since the WASL is not a test designed to establish a “ceiling for student cognitive abilities,” but rather one that seeks to set a minimum standard that all children are required to meet, any test questions pegged at the formal operational level are clearly inappropriate.

Orlich’s (2000a) analysis of the WASL showed it to be heavily weighted in the formal thinking area. Orlich (2000b) also illustrated how to use the scales of the National Assessment of Education Progress (NAEP) to determine cognitive reasoning in both standards to be met and test items. The combined inferences shown above lead us to posit that *Learned Helplessness*, *stress* and *depression* are unintended outcomes of the high-stakes test phenomenon. This has all the appearances of Sheila Tobias’ (1978) discussion of mathematics anxiety, which was closely aligned with learned helplessness. (See also Amrein and Berliner, 2002; Nichols and Berliner, 2005 for other unanticipated social consequences.)

Gunnar Myrdal, Swedish Nobel Laureate in Economics, published his seminal work--*An American Dilemma: The Negro Problem and Modern Democracy* (1944), which The Carnegie Corporation commissioned in 1938-40. Myrdal showed the ever-widening gap between equality and reality for African Americans in the United States. His work was subsequently cited

**Table 6. Percentage of Students at Piaget’s Cognitive Levels**

Age	Grade	Intuition	Entry Concrete (a)	Advanced Concrete (b)	Entry Formal (a)	Middle Formal (b)	Ref.
5.5	P	78	22				J
6	K	68	27	5			A
7	1	35	55	10			A, W
8	2	25	55	20			A
9	3	15	55	30			A
10	4	12	52	35	1		S
11	5	6	49	40	5		S
12	6-7	5	32	51	12		S
13	7-8	2	34	44	14	5	S
14	8-9	1	32	43	15	9	S
15	9-10	1	15	53	18	13	S
16	10-11	1	13	50	17	19	S
16-17	11-12	3	19	47	19	12	R
17-18	12	1	15	50	15	19	R
Adult	---	20	22	26	17	15	R

**Table Notes and References**

- Level (a) in each category is composed of children who have just begun to manifest one or two of that level’s reasoning schemes, while level (b) refers to children manifesting a half dozen or more reasoning schemes.
  - Table derived by Herman T. Epstein, personal communication with authors, June 8, 1999.
- J Smedslund, J. (1964). *Concrete Reasoning: A Study of Intellectual Development*. Lafayette, IN: Child Development Publications of the Society for Research in Child Development.
- A Arlin, P. Personal Communication with H. T. Epstein.
- W Wei, T. T. D., Lavatelli, C. B., & Jones, R. S. (1971). Piaget’s concept of classification: A comparative study of socially disadvantaged and middle-class young children. *Child Development*, 42 (3), 919-927.
- R Renner, J. W., Stafford, D. G., Lawson, A. E., McKinnon, J. W., Friot, F. E. & Kellogg, D. H. (1976). *Research, Teaching and Learning With the Piaget Model*. Norman: University of Oklahoma Press.
- S Shayer, M. & Adey, P. (1981). *Towards a Science of Science Teaching*. London: Heinemann.

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa “Summit on Public Education,” Washington, D. C.*

in the historic U. S. Supreme Court ruling in 1954, *Brown v Board of Education of Topeka*, which outlawed public school desegregation. Myrdal’s economic theory of cumulative causation posited that poverty breeds poverty.

The plethora of state-high-stakes-tests, in our opinion, have created **The New American Dilemma**. It has been abundantly shown with verifiable data that the poor, disfranchised, minority, and disabled children have fallen into education’s “achievement gap.” Perhaps the Carnegie Corporation might commission a 21<sup>st</sup> Century study to alert and sensitize policy-makers that a new dilemma now haunts our nation.

*In conclusion, the WASL in specific, and high-stakes tests in general, can be predicted to have deleterious effects on children of poverty, if not by simply instituting class-warfare and creating a permanent “underclass” with a hint of institutional racism.*

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa "Summit on Public Education," Washington, D. C.*

## References

Endnote: The authors acknowledge the critique of the statistics used by Dr. Marc A. Evans of Washington State University.

Abbott, M. L. & Joireman, J. (2001, July). *The Relationships Among Achievement, Low Income, and Ethnicity Across Six Groups of Washington State Students*. Lynnwood, WA: Washington School Research Center, Technical Report #1.

Adelman, C. (1999-2000). "The SAT Trap. Why Do We Make So Much of One 3-Hour Test?" *American Educator*, 23 (4), 38-41, Winter.

Amrein, A. L. & Berliner, D. C. (2002, March 28). "High-Stakes Testing, Uncertainty, and Student Learning." *Educational Policy Analysis Archives*, 10 (18), 1-56. Retrieved April 1, 2002 from <http://epasa.asu.edu/epaa/v10n18>.

Anderson, S. R. (1998, July). "The Trouble with Testing." *Young Children*, 53 (4), 25-29.

Beers, M. H. & Berkow, R. B. (Eds). (1999). *The Merck Manual of Diagnosis and Therapy*, 17<sup>th</sup> ed. Rahway, NJ: Merck, Sharpe and Dohme Research Laboratories.

Berliner, D. C. (2005). "Our Impoverished View of Educational Reform." *Teachers College Record*, August 2, 2005. Accessed at <http://www.tcrecord.org>.

Bolon, C. (2001, October 16). "Significance of Test-based Ratings for Metropolitan Boston Schools." *Education Policy Analysis Archives*, 9(42). Retrieved January 3, 2006, from: <http://asu.edu.epaa/v9n42/>.

CNN.com/EDUCATION. (2000, May 7). "Teachers Feel Stress on High-Stakes Testing." Retrieved June 18, 2002 from <http://fyi.cnn.com/2002/fyi/teachers.ednews/05/07/testing-pressure>.

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa "Summit on Public Education," Washington, D. C.*

Cohen, J. (1988). *Statistical Power for Behavioral Sciences*, 2<sup>nd</sup> ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cosgrove, J. (2000). *Breakdown: Facts About Stress in Teaching*. New York: Routledge Falmer.

Epstein, H. T. (2002). "Biopsychological Aspects of Memory and Education." In S. P. Shohov, Ed. *Advances in Psychology Research, Vol. 11*. New York: Nova Science Publisher, Inc., pp. 181-186.

*Fair Test Examiner*. (2004). Retrieved from [www.fairtest.org/nattest/ACT%20Scores%202004%20Chart.pdf](http://www.fairtest.org/nattest/ACT%20Scores%202004%20Chart.pdf). Retrieved from [www.fairtest.org/nattest/SAT%20Scores%202004%20Chart.pdf](http://www.fairtest.org/nattest/SAT%20Scores%202004%20Chart.pdf).

Firestone, W. A. & Mayrowetz, D. (2000). "Rethinking 'High States': Lessons from the United States and England and Wales." *Teachers College Record*, 102 (4), 724-749.

Fleming, J. & Garcia, N. (1998). "Are Standardized Tests Fair to African Americans? Predictive Validity of the SAT on Black and White Institutions." *Journal of Higher Education*, 69 (5), 471-495

Fouts, J. T. (2002, April). *The Power of Early Success: A Longitudinal Study of Student Performance on the Washington Assessment of Student Learning, 1998-2001*. Lynnwood, WA: Washington School Research Center, Research Report #1.

Gardner, J. W. (1961). *Excellence: Can We Be Equal and Excellent Too?* New York: Harper and Row.

Gottlieb, A. (2002). "Economically Segregated Schools Hurt Poor Kids, Study Shows. *The Term Paper*, 1 (2), 1-8.



*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa "Summit on Public Education," Washington, D. C.*

- Harrell, D. C. (2000, April 24). "Statewide Anxiety About WASL." *The Seattle Post-Intelligencer*, p. 1.
- Harington, M. (1962). *The Other America: Poverty in the United States*. Baltimore: Penguin Books.
- Harris, W. (2000, May 11). "WASL a 'Monster' of a Test." *The Spokesman-Review*, page 1-A.
- Hartstein, L. & Jones, A. (2002, April 12). "Bereaved Gwinnett Elementary School Has Many Unanswered Questions in Principal's Suicide." *The Atlanta Journal-Constitution*, retrieved on June 20, 2002 from <http://www.accessatlanta.com/ajc/metro/0402/12schooldeath.html>.
- Hattie, J. (1992). "Measuring the Effect of Schooling." *Australian Journal of Education* 36(1): 1-10.
- Hornbeck, M. (2001, March 11). "State Flunks 600 Schools." *Detroit News*. Retrieved on April 2, 2002 from <http://www.detnews.com>.
- Kids count DATA BOOK*. (2005). Baltimore: Annie E. Casey Foundaton.
- Knapp, M. S. & S. Woolverton. (2004). "Social Class and Schooling." In *Handbook of Research on Multicultural Education*, 2<sup>nd</sup> Edition. J. A. Banks and C. A. M. Banks, editors. San Francisco: Jossey-Bass, pp. 656-681.
- Lee, J. (2002). "Racial and Ethnic Achievement Gap Trends: Reversing the Progress Toward Equity?" *Educational Researcher*, 31 (1), 3-12.
- Lipman, P. (2004). *High Stakes Education: Inequality, Globalization, and School Reform*. New York: Routledge/Falmer.
- Marzano, R. J., Pickering, D. J. & Pollock, J. E. (2001). *Classroom Instruction That Works: Research-Based Strategies for Increasing Student Achievement*. Alexander, VA: Association for Supervision and Curriculum Development.

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa "Summit on Public Education," Washington, D. C.*

Marzano, R. J. (2004). *Building Background Knowledge for Academic Achievement.*

Alexander, VA: Association for Supervision and Curriculum Development.

Maslow, A. (1954). *Motivation and Personality.* New York: Harper.

McCallum, I. & Demie, F. (2001, Summer). "Social Class, Ethnicity and Educational Performance." *Educational Record*, 43 (2), 147-159.

Myrdal, G. (1944). *An American Dilemma: The Negro Problem and Modern Democracy.* New York: harper 7 Brothers, 2 vols.

Nairn, A. (Ed.). (1980). *The Reign of ETS: The Corporation That Makes Up Minds.* Washington, DC: Ralph Nader Report.

Nichols, S. & Berliner, D.C. (2005). *The Inevitable Corruption of Indicators and Educators Through High-Stakes Testing.* Tempe, AZ: Education Policy Research Unit, College of Education, Arizona State University. <http://www.greatlakescenter.org/pdf/EPSTL-0503-101-EPRU.pdf>.

Nyhan, R. C. & Alkadry, M. G. (1999, Fall). "The Impact of School Resources on Student Achievement Test Scores." *Journal of Education Finance*, 25 (2), 211-228.

Orlich, D. C. (2000). "A Critical Analysis of the Grade Four Washington Assessment of Student Learning." *Curriculum in Context*, 27 (2), 10-14. (a)

Orlich, D. C. (2000). "Education Reform and Limits to Student Achievement." *Phi Delta Kappan*, 81 (6), 468-472. (b)

Orlich, D. C. (2003). "An Examination of the Longitudinal Student Learning (WASL) on Student Achievement." *Education Policy Analysis Archives*, 11(18).

Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community.* New York: Simon & Shuster.

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa  
“ Summit on Public Education,” Washington, D. C.*

Sacks, P. (1999). *Standardized Minds: The High Price of America’s Testing Culture and What  
We Can Do to Change It.* New York: Perseus Books Group.

Sapolsky, R. (1994). *Why Zebras Don’t Get Ulcers.* New York: Freeman and Co.

Schoenfeld, A. H. (2002). “Making Mathematics work for All Children: Issues of Standards,  
Testing and Equity.” *Educational Researcher*, 31 (1), 13-25.

Shaw, L. (1999, October 10). “State Assessment Test Called Too Hard.” *The Seattle  
Times/Post Intelligencer*, pp. B1, B2.

Shim, M. K., Flener, R. D. & Shim, E. (2000, April). *The Effects of Family Structures on Aca-  
demic Achievement.* Paper presented at the Annual Meeting of the American Educational  
Research Association, New Orleans, LA.

Tobias, S. (1978). *Overcoming Math Anxiety.* New York: Norton.

UNICEF. “Child Poverty in Rich Countries, 2005.” *Innocenti Report Card No. 6.* UNICEF  
Innocenti Research Centre, Florence, Italy.

*This paper has been accepted for presentation at the October 20, 2006 Phi Delta Kappa  
“ Summit on Public Education,” Washington, D. C.*

### **The Authors**

Dr. Donald C. Orlich is Professor Emeritus, Science Mathematics Engineering Education Center at Washington State University (Pullman, Washington 99164-4237). His telephone number is (509) 335-4844 and e-mail address is [dorlich@wsu.edu](mailto:dorlich@wsu.edu). He is co-author of *Teaching Strategies: A Guide to Effective Instruction*, 8<sup>th</sup> Edition, (2007), Boston: Houghton Mifflin. In 2003, the University of Montana honored him with the “Educational Leadership Excellence Award.” He is author of *School Reform and The Great American Brain Robbery*, in press.

Dr. Glenn Gifford completed his Ph. D. in the Department of Educational Psychology at WSU. He specializes in educational psychology and applied psychometrics and is currently consulting in southern California.