

2008

## AYP Accountability Policy and Assessment Theory Conflicts

Randall S. Davies  
*Brigham Young University*

Follow this and additional works at: <https://scholarworks.bgsu.edu/mwer>

**How does access to this work benefit you? Let us know!**

---

### Recommended Citation

Davies, Randall S. (2008) "AYP Accountability Policy and Assessment Theory Conflicts," *Mid-Western Educational Researcher*. Vol. 21: Iss. 4, Article 2.

Available at: <https://scholarworks.bgsu.edu/mwer/vol21/iss4/2>

This Featured Article is brought to you for free and open access by the Journals at ScholarWorks@BGSU. It has been accepted for inclusion in Mid-Western Educational Researcher by an authorized editor of ScholarWorks@BGSU.

---

# *AYP Accountability Policy and Assessment Theory Conflicts*

Randall S. Davies  
Brigham Young University

## ***Abstract***

*A major objective of NCLB is to hold schools accountable for student achievement including closing the achievement gap and raising standards of student academic proficiency. While the majority of individuals support the underlying values of NCLB, not all agree its accountability policy is reasonable. The most problematic issue is the mandate that schools be held accountable to ensure all students perform at grade level when the methods used by states to determine grade level proficiency preclude the attainment of this standard. The unintended consequences of a policy that does not align theoretically with established educational best practice inevitably result in frustration of educational practitioners and often unwarranted condemnation of the educational system.*

The Elementary and Secondary Education Act (ESEA) of 2002 (Public Law 107-110), commonly referred to as No Child Left Behind (NCLB), represents an extensive reform of educational expectations and requirements. Of primary importance, NCLB made a substantial change in the rules that govern schools by requiring that they describe their success primarily in terms of student performance. According to the U.S. Department of Education (2002), a major objective of NCLB is to hold schools accountable for closing the achievement gap between various groups of students, with the goal that all students obtain a proficient level of achievement on state standardized tests.

NCLB holds educators accountable through a policy that requires each school to make adequate yearly progress (AYP) toward the goal of improving student achievement for all groups of students in their care. To measure AYP, the legislation mandates that states administer high quality annual assessments to every child at specified grade levels. These assessments must be aligned to standards consistent with nationally recognized professional and technical standards; they must be administered in a valid and reliable manner, and they must test higher order thinking skills using multiple measures (U.S. Department of Education, 2002). All public schools are expected to meet NCLB goals, testing, and reporting requirements. Schools receiving federal Title I funds are additionally subject to a series of corrective actions should they fail to meet AYP.

ESEA legislation has been in effect for over five years now, and while the majority of individuals support the underlying values of NCLB, not all agree that the law's implementation has had beneficial outcomes (Bracey, 2006; Center on Education Policy, 2006; Fuller, 2006; McElroy, 2005; Mathis, 2003; Weaver, 2007). At issue is the unintended consequences of establishing educational policy that is not consistent with educational assessment theory and best practice; more specifically, the policy mandating that schools be held accountable for student's achievement, and the unattainable expectation that all students perform at grade level on standardized tests (Linn, 2004; Raudenbush, 2004; Welner, 2005). While these expectations may seem reasonable, assessment methods used by states to determine grade level proficiency preclude the

attainment of this standard. The unintended consequence of a policy that does not align theoretically with established educational best practice understandably has resulted in frustration for educational practitioners and often unwarranted condemnation of the educational system in general.

## **Determining Proficiency and Academic Achievement Standards**

While NCLB mandates that all students should obtain a proficient level of achievement on state standardized tests, and has targeted the year 2014 for reaching this goal, the act allows each state to determine its own performance standards and definition of proficiency (U.S. Department of Education, 2004). As might be expected, the definitions and levels of proficiency differ widely from state to state (Rosenburg, 2004).

Performance standards, or student academic achievement standards as they are referred to in NCLB (P.L. 107-110, Section 1111(b)(1)(A)), are expected to be challenging; however, they must also be reasonable (Airasian & Russell, 2008; Linn, 2004). A requisite starting point for the process of establishing performance standards is a credible analysis of typical student performance in order to determine what can reasonably be expected from students at each grade level. Setting meaningful performance standards requires the use of established grade level norms, experience, and good judgment. It assumes that instruction at each grade level is based on a standardized curriculum that is aligned with state academic content standards. Establishing developmentally and educationally appropriate levels of performance is not always easy. Setting student achievement expectations too high or too low is undesirable in terms of motivation and obtaining increases in performance (Torrance, 1970). This is true for both teachers and students. There are political consequences to consider as well. Setting the standards too high will result in large numbers of students failing to pass state assessments. Setting standards too low invites criticism that students graduate without having to learn anything. Both situations are undesirable. Not only is setting performance standards a complicated process, it is clear that intelligent

---

people disagree on what the definition of proficient should be (Linn, 2003; Rosenburg, 2004).

State standardized tests are carefully created to assess student learning of content standards. This is important if the tests are to be valid estimates of student ability. Content standards tell teachers what will be tested and what they should teach. However, content standards do not indicate how well a student must perform in order to be considered proficient; performance standards are needed to make determinations regarding proficiency (Linn & Miller, 2005).

A common procedure for determining what a proficient grade level performance, or performance standard, should be is to have a group of educational experts and psychometricians make this decision (Airasian & Russell, 2008; Winchester, 2006). Based on their experience, these experts examine grade level norms, evaluate what students should be able to do at that age, consider appropriate expectations, and then set cut points for passing each exam. Cut points (i.e., passing scores) are typically reexamined on a regular basis to ensure they are appropriate, reasonable, and fair (Winchester, 2006). This method of determining standards for student performance adheres to accepted assessment theory and practices (Airasian & Russell, 2008; Linn & Miller, 2005; Reynolds, Livingston & Willson, 2006). What this means is that not everyone who takes the test will pass; it is an impossibility when the passing score is based on an analysis of typical performance.

### *Criterion and Norm Referenced Assessments*

A common error in understanding for many people suggests that because state standardized tests are supposed to be criterion referenced, not norm referenced, the use of a cut point to indicate proficiency does not affect the number of students who can achieve proficiency. If standardized tests were designed to test mastery, this may be true; but in practice, this assumption is incorrect on several accounts. Most importantly, the concept of criterion reference refers to how the results are interpreted and reported, not how a proficiency cut point is determined, and thus, how many students would reasonably be expected to pass the test (Airasian & Russell, 2008; Linn & Miller, 2005; Reynolds, Livingston & Willson, 2006).

State standardized tests are considered to be criterion referenced due to the fact that individual results are interpreted against content standards rather than the performance of other students taking the exam each year. Unfortunately, state tests are typically not designed to maximize this type of interpretation. State standardized tests typically cover a large domain of learning tasks across multiple standards which limits the number of questions that can be asked. Test items are selected based on their discriminating power and typically exclude items that are too easy or too difficult (Winchester, 2006). This type of test design is better suited to norm referenced interpretations (Linn & Miller, 2005); however, state tests are designed in this way to accommodate item response theory requirements (Indiana Department of

Education, 2007). In order to make use of item response theory, test makers routinely and justifiably use typical performance analysis and grade level norms to determine item difficulty and discriminating power rather than using absolute measures of content mastery. As a result, even though the interpretation and reporting of these test results are criterion referenced, acceptable individual student performance is determined by a comparison of typical student performance based on grade level norms expectations.

This is not a criticism of state examinations. Based on the evidence, these tests seem to provide valid and reliable estimates of student ability. At issue is how the results are used (i.e., consequential validity); more specifically, while one might hope that all students in a particular grade would be able to master the curricular content, because passing scores are based on an analysis of typical student performance, it is unreasonable to expect all students would pass state standardized tests. Given the way passing scores are determined, some students will always be below grade level on state standardized tests, and the number of students below grade level will be fairly constant. Because AYP is based on the percentage of students passing state standardized tests—and it is impossible for all students to pass the exam given the method for determining passing scores—inevitably all schools will eventually fail to meet increasingly more challenging AYP benchmarks.

### *Expected Results When Using Typical Performance to Determine Proficiency*

In theory, the practice of establishing appropriate grade level performance standards is based on the belief that some human characteristics (e.g., intelligence, cognitive ability, and academic achievement) are normally distributed in the population. There will always be a few individuals who are considerably more able or skilled than the average; and a few individuals who will always be significantly less able or skilled compared to the average; however, a large percentage of students (68%) will demonstrate average performance (i.e., one standard deviation above or below the mean). Statistically, students who score in this range are arguably quite similar given the fact that all tests are prone to a certain amount of measurement error and an individual's true scores must be considered within a derived confidence interval (Linn & Miller, 2005). Still, a cut point must be established, and inevitably it will exclude some students from passing the exam.

Students who pass state tests are deemed proficient; students who fall significantly below the range of average performance established for a specific assessment are by definition considered to be below grade level. The concept of grade level proficiency suggests a band of acceptable performance based on the observed typical performance of students in that grade (American Federation of Teacher, 2004; Linn & Miller, 2005). If proficiency is defined as those students who obtain a passing score on a test, and the passing score

is determined through an analysis of typical performance for students in that grade, then it is only reasonable to expect that a group of students will always be excluded from being labeled proficient. The exact percentage of students who fail to reach a proficient level of achievement will depend on where the cut point is set.

Obviously, if the passing score of a test is set to equal the mean or average score obtained by students taking the test, by definition half of the students (50%) would pass and the other half would fail. It would be unreasonable to expect that all students pass the test if this were the case, because by definition it is impossible for more than 50% of the students taking any given test to be above average. However, a very similar thing happens when a cut point is used to determine proficiency instead of the mean.

An illustration of this phenomenon is presented in Table 1. Since all states use different assessments and set different cut points, it is more appropriate to look at each state individually. For the purposes of this paper, the state of Indiana is used as a case study. Table 1 data represent pass rates for the third grade Language Arts and Mathematics assessment in the state of Indiana over the past five years; similar performance patterns can be observed in other states and for other grade levels.

Table 1  
*Statewide 3<sup>rd</sup> Grade Language Arts Assessment Results Comparison by Year*

Year	Mean	SD	Z score of Passing Score	% Passing
2002	436.7	61.8	-0.53	72
2003	442.4	65.5	-0.59	74
2004	443.0	63.5	-0.61	75
2005	442.6	64.9	-0.59	75
2006	438.9	67.3	-0.56	74
2007	442.1	64.2	-0.60	75
Average	441		-0.58	74

*Statewide 3<sup>rd</sup> Grade Mathematics Assessment Results Comparison by Year*

Year	Mean	SD	Z score of Passing Score	% Passing
2002	416.0	61.2	-0.38	67
2003	424.0	63.7	-0.49	71
2004	426.5	62.7	-0.53	73
2005	426.1	64.8	-0.51	73
2006	425.0	65.8	-0.49	72
2007	420.5	64.2	-0.43	70
Average	423		-0.47	71

Established cut points for each year was 393 for Math and 404 for Language Arts

As assessment theory predicts, in both situations you get a consistent percentage of students passing the exam each year. Variations in the result are just as likely explained by differences in the exams given each year, teachers better aligning their instruction with state content standards, or by

schools better preparing students to take exams as they are by any real change in the ability of the students from year to year. Analysis of scores from the National Assessment of Educational Progress (NAEP) confirm this result suggesting no significant overall improvement in student achievement has occurred since the enactment of NCLB legislation (Bracey, 2006; Center on Education Policy, 2006). Given that in theory achievement and scholastic ability are believed to be normally distributed and that the criteria for establishing specific proficiency cut points are based on an analysis of typical performance, this result would be expected (Armor, 2006; Reynolds, Livingston & Willson, 2006). If anything, this situation suggests that the assessments used are somewhat reliable at measuring what they measure.

In general, achievement trends in each state show the overall percentage of students that pass state standardized assessments each year has been, and will likely remain, fairly constant for students as an overall group. One thing is clear; equating school quality with the percentage of students at that school who achieve “proficiency” does not withstand serious scientific scrutiny when the standard for proficiency is based on an expectation of typical grade level performance (Raudenbush, 2004). The assumption that schools can somehow cause all students to perform at or above grade level proficiency constitutes an unrealistic expectation (Linn, 2004; Welner, 2005).

## Breakdown of Results

There are many reasons why children are left behind in our schools; lacking the opportunity to receive a high-quality education is but one (Reigeluth & Beatty, 2003).

A well-substantiated body of research links teacher quality and student learning (Darling-Hammond, 2000; Mendro, 1998; Stedman, 1997; Wenglinsky, 2002), with a particularly strong positive correlation between teacher subject area certification and student achievement (Goldhaber & Brewer, 1996). However, being highly qualified and providing a quality learning opportunity does not mean teachers will be completely effective for all students in all situations. Meeting AYP targets is not a valid indicator of teacher and school quality; there are better explanations for understanding poor student performance in schools (Raudenbush, 2004).

Continuing with the Indiana example, consider a comparison of three school districts. In Indiana, school districts are called school corporations. Table 2 shows a demographic comparison for three school corporations all located in one metropolitan area which includes the surrounding rural population. These school corporations are largely divided along lines of Socio-Economic Status (SES) with larger minority populations and a higher proportion of special needs students in the low SES areas. School Corporation A is low achieving, Corporation B has moderate achievement, and Corporation C is high achieving (see Table 3). School Corporation A failed to meet their AYP targets in the 2006-2007 school year. Both the other school corporations met AYP goals that year. These



three school corporations are quite diverse, yet together they constitute a reasonably representative sample of the state. Although the corporations combined have slightly more minorities, more students receiving free or reduced lunch, and more students with special needs, they correspondingly have a slightly lower percentage of students passing the state assessments in language arts and mathematics. While there is some variation in the overall result, as expected, the average number of students in the state passing the test and the average number of students in these combined school corporations passing the test has varied little over the past five years (see Figures 1 and 2).

All teachers in the state of Indiana are required by law to be highly qualified to teach in their subject areas, yet these three corporations have consistently obtained very different results. On the surface, based on the percentage of students passing state tests, School Corporations A and B might be seen as less effective. Some might erroneously believe that the teachers in these schools are less qualified, dedicated, or able. However, looking at the trend analysis in Figures 1 and 2, these two school corporations had the greatest gains in achievement over the past five years.

One important aspect related to school success is the portion of special needs students being served by a school. Schools with large proportions of special needs students are typically at greatest risk of failing to meet AYP expectations (Kim & Sunderman, 2005). In Indiana during the 2006-2007 school year,

44 (71%) of the 62 school corporations that failed to meet AYP targets failed in the special needs category. Thirty (48%) of the 62 corporations failed only in this area. This is not surprising given that special need students understandably do less well on tests than their more able peers. When compared to regular students, special needs students will always end up at the low end of the assessment distribution.

Another pattern in these data is the correlation between the number of students receiving Free or Reduced Lunch and academic achievement. Again, schools with large proportions of students living in poverty are at greater risk of failing to meet AYP expectations (Marzano, Pickering, & Pollock, 2005; Holmes-Smith, 2006; Heck, 2006). Notably, minority students are more likely to be living in poverty; however, this result might suggest that it is not a students' minority status that influences scholastic performance, but rather the fact that they live in poverty and are more likely to have special learning needs. Teaching students under these circumstances represents a special challenge for schools (Heck, 2006).

Given this pattern of achievement, none of these school corporations are on track to have all their students at a proficient level of achievement on state standardized tests by the 2013-2014 school year. The only logical conclusion one can make is that each of these school corporations will eventually and inevitably be labeled as "failing" (Armor, 2006). Regardless of the reasons students fail to meet grade level expectations, eventually all schools will fail to meet the ever

Table 2  
*School Corporation and State Demographics for 2006-2007*

	Total Enrollment	% Minorities	% Free/Reduced Lunch	% Special Needs	Graduation Rate	Drop Out Rate
Corporation A	21874	68	62	24.6	60.4	22.8
Corporation B	5682	22	48	20.4	52.6	18.6
Corporation C	10612	13	15	14.6	83.2	15.0
Combined	38168	46	47	21.2	65.6	20.0
State	1045702	33	38	17.8	76.5	11.2

Table 3  
*Percent Passing Language Arts Assessment by School Corporation and Year*

	2002	2003	2004	2005	2006
Corporation A	51.3	50.8	55.0	54.3	56.8
Corporation B	63.6	62.4	65.9	67.1	69.6
Corporation C	81.6	82.7	82.7	83.1	84.4
Combined	61.2	61.4	64.2	64.1	66.4
State	59.4	70.6	71.2	71.7	71.4

*Percent Passing Mathematics Assessment by School Corporation and Year*

	2002	2003	2004	2005	2006
Corporation A	51.8	53.0	56.4	57.5	58.5
Corporation B	68.3	70.3	72.3	75.1	75.5
Corporation C	83.7	85.3	85.3	86.2	85.7
Combined	62.8	64.6	66.7	68.0	68.6
State	68.0	71.4	72.2	74.0	73.7

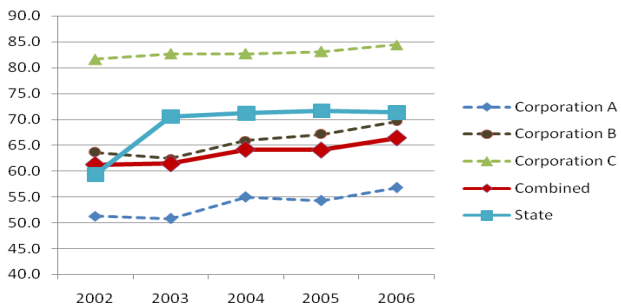


Figure 1. Percent Passing Language Arts Assessment by School Corporation and Year

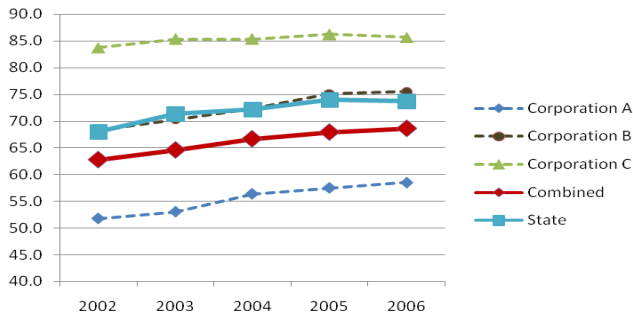


Figure 2. Percent Passing Mathematics Assessment by School Corporation and Year

increasing targets of AYP. Because of the way proficiency is defined and measured, even schools recognized for excellence will eventually succumb.

### Inevitable and Unanticipated Consequence of Impossible Expectations

Making adequate yearly progress (AYP) is primarily defined as the degree to which schools meet specific targets in the number of students at or above proficiency on state examinations (American Federation of Teacher, 2004). This applies to all students at the school, including disaggregated groups of students within a school. In order to attain a positive AYP status and to avoid the “failing school” label, many schools have at times resorted to exceptional practices in an attempt to maximize student achievement on NCLB mandated standardized tests (Orina & Davies, 2006). Teachers often note feeling relieved once standardized testing is done each year so they could get back to the normal tasks of educating children. One unintended consequence that is becoming harder for teachers to overcome is the stress, frustration, and anger they feel at being held accountable for things over which they have little or no control (Orina & Davies, 2006; Sulok, 2005).

Regardless of assurances that teachers would not be held individually responsible for a school’s failure to achieve AYP, many teachers feel they are being blamed for low test results in spite of their best efforts. Parents are understandably concerned when a school fails to meet AYP. The National Center for Educational Statistics (2007) estimates that approximately

27% of schools failed to meet AYP in the 2004-2005 school year nationwide. It is expected that an increasing number of schools will fail to meet AYP targets each year (Wiley, Mathis, & Garcia, 2005). As this happens, teacher morale can be expected to decrease even more.

This is especially true for teachers who work with students in various special needs categories. Often these students are held to the same performance standards as general education students even though they will eventually end up at the low end of the achievement distribution. It is clearly not reasonable to blame special needs students for this situation, nor is it reasonable to blame schools when they cannot reach unrealistic expectations of performance.

In many schools the special needs populations are increasing (Gunter, 2005). Schools with large populations of students in these categories often experience an enormous drain on resources attempting to provide for these students and ensure they achieve. In fact, students who do not perform adequately often become the main focus of instructional efforts to the exclusion of those groups of students already attaining proficient levels of achievement (Orina & Davies, 2006). Teachers in schools with large numbers of students identified in ENL or special needs categories often indicate a sense of futility at meeting the testing requirements (Orina & Davies, 2006). They are becoming overwhelmed with the challenge of making sure these students test well in comparison to the general education student population. Critics of NCLB argue that federal efforts demanding results-based accountability are presumptively futile because they assume that all students can and will perform at a proficient level academically regardless of their abilities and motivation (Schrag, 2004). One might argue that teacher morale is declining as a direct result of the pressure being put on them to meet unrealistic accountability standards that are becoming more and more unattainable (Sulok, 2005). In many ways the mandate to ensure all students achieve proficiency is a “shoot for the moon; even if you miss, you’ll land among the stars” philosophy. The fact that teachers and schools are being asked to meet unrealistic AYP goals does not seem to matter.

### Discussion Summary

In general, the methods for determining grade level proficiency used by most states are appropriate and in accordance with accepted assessment theory. Setting meaningful performance standards based on grade level proficiency requires the use of established grade level norms and good judgment. Certainly there is disagreement between states on what it means to set challenging student academic achievement standards, but this is to be expected given the nature of human beings and the often conflicting and varied views regarding the purposes of public education.

The most problematic aspect of NCLB policy is that the way states define grade level proficiency will always exclude some students from being categorized as proficient. This would not be a problem except for the fact that by equating

---

the cut point (i.e., passing score) on state standardized test with performance standards, and then expecting all students to be able to meet that standard, constitutes an unrealistic expectation and impossible burden for the educational system. Certainly some schools currently enjoy a greater level of success in terms of the number of students passing state assessments compared to other schools; however, all schools, even the best schools, will eventually reach an achievement barrier in which they cannot do any better in terms of meeting ever increasing AYP targets.

While most teachers feel it is acceptable to expect that their students attain an adequate level of achievement, there is a growing feeling of frustration among educators. Often this frustration seems to stem from the fact that teachers are being held accountable for factors and conditions they have no ability to change. Important factors that affect learning include: regular attendance at school; the effort and attention students put into their studies; the support and encouragement students receive from home; the economic and societal influences they experience; and the students' abilities, willingness, interests, and intentions for the learning expected of them as students. These and other issues are seen by teachers as factors they cannot control. Unfortunately, AYP status is considered by some to be synonymous with school quality. State assessments of student ability are not valid indicators of school and teacher quality. Despite the best efforts of excellent teachers, many students do not attain proficient levels of achievement, and even in the best classrooms, not all students achieve excellence.

### Conclusions

Assessment is a fundamental aspect of the teaching and learning process; however, a proper understanding of assessment is essential when establishing educational policy. Policy that does not align theoretically with established educational best practice inevitably results in frustration of educational practitioners and often unwarranted condemnation of the educational system. One aspect of the current ESEA's accountability policy that does not align with educational theory is the expectation that all students can meet grade level standards when such standards of proficiency are based on the normal distribution of typical student performance for that grade level. Federal mandates that expect schools to meet arbitrary AYP targets should be removed from NCLB accountability policy. Students should be tested, but it is inappropriate to use results of state standardized assessments as the primary evidence for judging the quality of teachers and schools.

NCLB seems to operate on the premise that every child can learn and will learn if they are provided with a highly qualified teacher and a beneficial learning environment that utilizes scientifically proven practices. However, there are many reasons why children are left behind in our schools; lacking the opportunity to receive a high-quality education is but one (Reigeluth & Beatty, 2003). Educational accountability policy needs to hold teachers and schools accountable

for things they have control over. It is the responsibility of educators to provide quality instruction within a beneficial learning environment. They should care about their students and do what they can, within reason, to help them learn. It is the students' responsibility to take advantage of the opportunities they are given. Students should be held accountable for their own achievement. Assessment results help teachers and students identify areas for improvement. While there are likely several things schools and individual teachers can do to improve instructional practices, meeting unrealistic targets for student achievement when those targets are by definition impossible to reach is something schools should not be burdened with.

### References

- American Federation of Teacher (2004). *Eight misconceptions about the No Child Left Behind act's (NCLB) adequate yearly progress (AYP) provisions*. Retrieved November 30, 2006, from <http://www.aft.org/topics/nclb/downloads/8Misconceptions.pdf>
- Airasian, P. W., & Russell, M. K. (2008). *Classroom assessment: Concepts and applications* (6<sup>th</sup> ed.). Boston: McGraw-Hill.
- Armor, D. J. (2006, August). Can NCLB close the achievement gap? *Teachers College Record* (ID Number: 12667). Retrieved August 23, 2006, <http://www.tcrecord.edu.org>
- Bracey, G. W. (2006). The 16<sup>th</sup> Bracey report on the condition of public education. *Phi Delta Kappan*, 88(2), 151-166.
- Center on Education Policy (2006). *From the capitol to the classroom: Year 4 of the No Child Behind Act, summary and recommendations*. Retrieved November 23, 2006, from <http://www.cep-dc.org/nclb/Year4/NCLB-Year4Summary.pdf>
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives*, 8(1). Retrieved March 30, 2007, from <http://epaa.asu.edu/epaa/v8n1>
- Fuller, B. (2006). *Is the No Child Left Behind act working?* Berkeley, CA: Policy Analysis for California Education.
- Goldhaber, D. D., & Brewer, D. J. (1996). Evaluating the effect of teacher degree level on educational performance. *ERIC Digest*. (ERIC Document Reproduction Service No. ED406400).
- Gunter, C. (2005). *Educating students with limited English proficiency*. Unpublished masters project. Indiana University, South Bend.
- Heck, R. H. (2006). Assessing school achievement progress: Comparing alternative approaches. *Educational Administration Quarterly*, 42, 667-698.
- Holmes-Smith, P. (2006). *School socio-economic density and its effect on school performance*. Retrieved October, 21, 2007, from [http://www.mceetya.edu.au/verve/\\_resources/SES\\_Report.pdf](http://www.mceetya.edu.au/verve/_resources/SES_Report.pdf)
- Indiana Department of Education. (2007). *Guide to test interpretation*. Retrieved October 30, 2007, from [http://www.doe.state.in.us/istep/pdf/GTI/49640-W\\_GTI\\_s07IN.pdf](http://www.doe.state.in.us/istep/pdf/GTI/49640-W_GTI_s07IN.pdf)

- Kim, J. S., & Sunderman, G. L. (2005). Measuring academic proficiency under the No Child Left Behind act: Implications for educational equity. *Educational Researcher*, 34(8), 3-13.
- Linn, R. L. (2003). Accountability: Responsibility and reasonable expectations. *Educational Research*, 31(7), 3-13.
- Linn, R. L. (2004). *Rethinking the No Child Left Behind accountability system*. Paper presented at a forum sponsored by the Center on Education Policy, Washington, DC.
- Linn, R. L., & Miller, M. D. (2005). *Measurement and assessment in teaching* (9<sup>th</sup> ed.). Upper Saddle River, NJ: Pearson Education.
- Mathis, W. J. (2003). *No Child Left Behind: Costs and benefits*. Phi Delta Kappa. Retrieved on August 7, 2006, from <http://www.pdkintl.org/kappan/k0305mat.htm>
- Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2005). *Classroom instruction that works*. Upper Saddle River, NJ: Pearson Education.
- McElroy, E. J. (2005, January). *Teaching left behind. Teaching Pre K-8*. Retrieved August 7, 2006, from [http://www.teachingk8.com/archives/teacher\\_to\\_teacher/teaching\\_left\\_behind\\_by\\_edward\\_j\\_mcelroy.html](http://www.teachingk8.com/archives/teacher_to_teacher/teaching_left_behind_by_edward_j_mcelroy.html)
- Mendro, R. L. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. *Journal of Personal Evaluation in Education*, 12, 247-256.
- National Center for Educational Statistics (2007). State educational reforms: Data tables. Retrieved October 10, 2008, from [http://nces.ed.gov/programs/statereform/saa\\_tab6.asp](http://nces.ed.gov/programs/statereform/saa_tab6.asp)
- No Child Left Behind Act of 2001, Pub. Law No. 107-110, 115 Stat. 1425 (2002).
- Orina, E., & Davies, R. (2006, Mar). *The unintended consequences of NCLB*. Presented at the annual conference of the Michigan Academy of Science, Arts, and Letters (MASAL), Rochester, MI.
- Raudenbush, S. W. (2004). Schooling, statistics, and poverty: Can we measure school improvement? *Educational Testing Service, William H. Angoff Memorial Lecture Series*. 6.
- Reigeluth, C., & Beatty, B. (2003). Why children are left behind and what we can do about it. *Educational Technology*, 43(5), 24-32.
- Reynolds, C. R., Livingston, R. B., & Willson, V. (2006). *Measurement and Assessment in Education*. Boston: Allyn and Bacon.
- Rosenburg, B. (2004). *What's proficient? The No Child Left Behind act and the many meanings of proficiency*. Retrieved November 30, 2006, from <http://aft.org/pubs-reports/downloads/teachers/WhatsProficient.pdf>
- Stedman, L. C. (1997). International achievement differences: An assessment of a new perspective. *Educational Researcher*, 26, 4-15.
- Shrag, P. (2004). *Bush's education fraud. The American prospect*. Retrieved on August 7, 2006 from <http://www.prospect.org/web/page.wv?section=root&name=ViewPrint&articleId=6998>
- Sulok, N. (2005, December 12). MEA President says "No Child" hurting morale. *South Bend Tribune*. Retrieved on November 30, 2006, from <http://www.southbendtribune.com/apps/pbcs.dll/article?AID=/20051212/News01/512120316/0/BIZ>
- Torrance, E. P. (1970). *Encouraging creativity in the classroom*. Dubuque, IA: Wm. C. Brown.
- U.S. Department of Education, Office of Elementary and Secondary Education (2002). *No Child Left Behind: A desktop reference*, Washington, DC.
- U.S. Department of Education (2004). *A guide to education and No Child Left Behind*. (Office of the Secretary, Office of Public Affairs). Washington, DC.
- Weaver, R. (2007). NCLB: The view up close. *Social Policy*, 37(3/4), 45-48.
- Welner, K. G. (2005). Can irrational become unconstitutional? NCLB's 100% presuppositions. *Equity and Excellence in Education*, 38, 171-179.
- Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives*, 10(12). Retrieved June 20, 2008, from <http://epaa.asu.edu/epaa/v10n12/>
- Winchester, D. (2006, October). *Indiana academic standards and benchmarks*. Educational technology meeting presentation, Indianapolis, IN.
- Wiley, E., Mathis, W., & Garcia, D. (2005). *The impact of the adequate yearly progress requirement of the federal "No Child Left Behind" act on schools in the Great Lakes region*. Retrieved November 30, 2006, from [http://greatlakescenter.org/g\\_l\\_new\\_doc/EPSE-0505-109-EPRU\\_EXEC.pdf](http://greatlakescenter.org/g_l_new_doc/EPSE-0505-109-EPRU_EXEC.pdf)