

History, Validity and Washbacks of NCLB Testing

Ustun Kaan

Abstract

This final research paper is mainly intended for the Second Language Testing students, those interested in teacher engagement and in language teaching. This final research paper will focus briefly on the history, the implementation, the validity as well as reliability and various washbacks of the NCLB (No Child Left Behind) law signed by President George W. Bush in 2002. This research will also detail some of the struggles teachers face such as engagement, motivation, and even at often times, frustration and pressure due to yearly test results required by the NCLB law and its influence on their teaching methods. I based my information on online articles, online websites such as the *Accountability Resource Center in Hawaii*, previous observations and experiences in a Hawaiian public school classroom, some literature review as well as in class theory. Through this final research paper I will attempt to answer some questions related to the NCLB program. Different literatures pointed me to many questions about that program such as its validity due to various influencing factors ( money, politics, teacher quality, cost of testing, fund distribution,...), the potential consequences (washbacks) of the program on the students' learning process, as well as the crucial question of accountability. What is NCLB ? How is teacher engagement defined? What are the criterias of teacher engagement? How are those criterias related to NCLB? And finally, what are some of the washbacks, positive and negative, from the NCLB program? Finally, I will conclude this paper by sharing some of my thoughts on the validity of the program, its reliability and how my previous experience in a Hawaiian public school classroom influenced my position on the topic of engagement, the NCLB law as well as my own philosophy of teaching.

## Introduction

The literature review was driven by three different points: a historical overview of the NCLB Testing, the issue of accountability through the various assessments and problems related to the implementation of the assessments themselves. However, at the heart of NCLB (and now Race To The Top) lies one true question: who is accountable? In order to determine who is accountable for test results, a test must first be confirmed to be valid, reliable and fair, and the teaching must be adequate and of high quality. There is another crucial factor to consider in this equation : classroom composition, or in other words, the students! “One of the most critical challenges facing educators today in the USA is the ever-increasing enrolment of students in all grades K-12 for whom English is a second language”(Butler; Stevens, 2001). Furthermore, “depending on a multitude of factors – such as age of arrival and prior formal education – many children take a number of years to master the language of education of the USA, academic English, before they can participate and learn equally alongside native English speakers” (Cummins,1981;Collier,1989; Hakuta et al.2000) (Butler; Stevens, 2001). I often wondered who truly was responsible for test results, the classroom education and ultimately the nations’ education level and standing in the world. Is it the teachers? Is the parents? Or is it the system of education in place today ? Throughout the semester we covered all aspects of testing, from its validity to its reliability, from developing various good quality language test items to describing language test results, interpreting them and finally the correlation between different language test results and their meaning. One of the traditional definition of validity provided in class was “discovering whether a test measures accurately what it is intended to measure” (Hughes, 1989). So, are the NCLB assessments reliable? Are they valid? According to Butler and Stevens, “the test scores are not always reliable and valid for all the purposes for which they are used or for all the students to whom they are given” (2001). The reasons Butler and Stevens invoke to support

their claim is that “commercially developed large-scale content assessments were produced and intended for students who are native speakers of English or highly proficient non-native speakers” (Davidson,1994) (Butler; Stevens, 2001). Furthermore, they go on to state that “at the present time (2001) however, across states, large numbers of ELL are expected to take standardized assessments with their native English-speaking peers, regardless of how long they have lived in the USA or how well they have mastered English “ (Holmes and Duron, 2000) (Butler; Stevens, 2001). However, as I read further, the case against the reliability of NCLB did not seem as clear a black and white issue as it seemed to be 10 years ago. Did NCLB evolve ? Is it more fair today? Or has it become worse? I feel a brief historical overview of the NCLB program is necessary to understand the questions at hand. NCLB is a deeply complex issue that despite having positive goals at its heart such as encouraging as many students as possible to become more proficient in reading, science, math and as a consequence, in the English language, enhancing the education level in the USA, holding bad teachers accountable for their bad teaching methods, and ultimately sending more students to college, it seems flawed in its implementation and results. But how did it all start?

### Past, Present and Future

It is interesting to notice that “NCLB is a direct descendent of the Elementary and Secondary Education Act of 1965, a key component of the legislation that identified President Lyndon Johnson’s Great Society in the 1960s” (Bunch, 2011). Bunch goes on to state that “it is noteworthy that the 1965 legislation mentioned neither ELL students nor educational accountability for districts and states” (2011). It is 2 years later that “congress recognized a special class of students – those with limited English proficiency – in addition to the children of

poverty, for which the law was originally designed” (Bunch, 2011). In 1983, “a panel of nationally recognized experts published *A Nation at Risk*, noting, among other things, that there were too few teachers qualified to teach ELL children” (Bunch, 2011). Today, teachers are the first in the line of duty of education. Their quality is a topic of much debate. The stakes are high, the needs of the students keep growing, and the pressure on the teachers ‘shoulders are stronger than ever. As for the assessment criterias, “much of the language in NCLB pertaining to assessment of ELL students is based on (...) reforms”(Bunch, 2011), implemented by NAEP (National Assessment of Educational Progress) since 1964. Some of those reforms include “development of assessments based on clearly defined content standards; regular, periodic collection of assessment data, assessments characterized by technical rigor” (Bunch, 2011). Finally, as Bunch explains, “NCLB reflects the many changes in our nations’ view of public education – who should receive its benefits and how we should hold educators accountable – over the past 45 years” (2011). In other words, where does the blame of our education system lie in times of crisis, which we seemed to find ourselves in the midst of right now. How can we fix it, and would the results be reliable? People often use the line: it all starts in the classroom. The it in question is our childrens’ education, in essence, their future. Is it the governments’ responsibility to make sure its citizens are well prepared for the 21<sup>st</sup> century world? Today, according to Craig Deville and Micheline Chalhoub-Deville,

“the government of President Barack Obama continues the push of the previous administrations for a reform agenda based on accountability assessment system that includes national standards and assessments” (2011)

The US Department of Education claims that this agenda will challenge

“ the nation to embrace education standards that would put America on a path to global leadership [and] provide incentives for states to adopt academic standards that prepare students to succeed in college and the workplace, and create accountability systems that measure student growth toward meeting the goal that all children graduate from high school and succeed in college”(US Department of Education, 2010)

But will it? There are times when the government needs to step in, as the adult in the room, to make sure its children are protected, educated and have a chance at a better future. But is this future defined by grades on a piece of paper? If so, shouldn't everyone be involved in the process to make sure it is reliable and valid? This beautiful ideology, this ideal philosophy of a better future sadly does not hold in front of the realities of our world today. A famous coach once said that “there is no I in team”. “The national content standards, labeled CCS (Common Core Standards) are a critical component of the administrations' future vision and plans” (Deville; Chalhoub-Deville, 2011). However, as I read further, I was saddened to learn that the process by which standards of assessments were created, as well as the curriculum, did not include the most important people: the ones in the education system.

“Mathis adds that professionals working in the schools were excluded from this process:”Practitioners and subject matter experts complained that they were excluded from the development process” (p.5). Essentially, the process used to develop the CCS represents a top-down approach that deliberately excludes education professionals and practitioners who are the people who will be held responsible for the implementation and success of the system” (Deville;Chalhoub-Deville, 2011)

The future seems dark. Messick's validity theory states that "validity is not a property of a test or assessment but the degree to which we are justified in making an inference to a construct from a test score". How valid can a test score be if the people at the heart of them are not included in the process of creating them? It is *the do as I say and do it good or you will get fired type rhetoric*. It leads me to the following quote we were given in class:

"All examinations are symbol of power and control" (Foucault, 1975).

Is NCLB about control? Is there a conspiracy theory going on? Some people seem to think so.

"While a nationwide survey of public school administrators (Archer,2003) indicates broad acceptance to raise standards for student performance, there is widespread skepticism over the law's workability. About half of those polled believe the federal legislation was either politically motivated or aimed at undermining public schools. Rose (2003) protests that NCLB's accountability requirements make it a 'rigged game' and a 'statistical trap' for public schools (p.1). Gerald Bracey labeled the act 'a weapon of mass destruction and its target is the public schools' (Hill; Barth, 2004)

This rhetoric seems to spring out of anger and frustration of some of the washbacks of the NCLB legislation. One the strongest consequences concerns the teachers, their qualifications, teaching methods and ultimately their retention by the numerous public schools. Let me turn now to the teacher engagement, qualification and retention problems, and other washbacks.

#### Engagement, qualifications, washbacks and how to retain our teachers

Teacher engagement is crucial to the well being of a class, a school, a state and finally a country, in any part of the world. It all starts with one engaged teacher willing and eager to teach.

Teacher engagement includes many different lives. There are happy lives, respectful ones, sad ones and others who are just between life and death, barely surviving. The many aspects of teacher engagement are linked to various and numerous factors within the classrooms, the school settings, and the education system. Engagement in a classroom is a two way street between the students and the teacher. Teacher engagement within the school settings and education system relies on and is a result of accessible benchmarks , teaching methods, available teaching materials, classroom sizes and environment. “The debate then, centers on the question of whom and what defines teacher quality” (Hill; Barth, 2004). Furthermore, “raising teacher quality has become education reform’s top priority, and upholds that ‘teaching quality’ is the single most important factor influencing student achievement” (Kaplan & Owings, 2002, p.21) (Hill; Barth, 2004). Many factors create confusion over the role and responsibility of the teacher in a classroom. The how, the why, the what to teach all contribute to the crucial feeling of engagement which must be shown by teachers. The whom is another important factor: students with learning disabilities , students part of the gifted and talented program, or English Language Learners students from multiple cultural and socio economic backgrounds, all require different types and levels of engagement. Finally, the economic ingredient plays a major part in getting teachers engaged. The economics such as public school funding to provide better classroom environments and teaching materials is linked to benchmarks. Benchmarks in reading, writing, math or science seem to have become the holy grail of teachers in the United States education system. Nonetheless, those benchmarks directly influence the teaching methods, and the engagement in teaching them. Each classroom, each teacher, each subject has its own life, its own teaching methods and its own benchmarks. My focus here is on ELL teachers.



According to the Blueprint for reform from the current administration:

“The Blueprint is based on two principles we all know are true: 1) Great teachers matter and 2) Not all teachers are equally effective. Under this plan, teachers are recognized and supported as unique professionals. Principals and other school leaders will take the time (and they should be given the training) to evaluate teachers comprehensively and fairly based on individual performance. The plan proposes that evaluations be based on multiple measures, acknowledge success, provide meaningful feedback, inform staff development and staffing decisions, offer teachers more responsibility, and compensate them more reasonably. Teachers deserve to be evaluated fairly and paid for the hard work they do.” (US Department of Education, 2010)

It is a clear change from when the former secretary of education Rod Paige stated that “we can only measure a teacher’s success through the improvement of his or her students; in my opinion, there is no other measure (Renwick, 2002, p.24) (Hill; Barth, 2004). However, today’s message sounds like a true message of support and encouragement for the ones on the front lines.

Nevertheless, if teachers are to be evaluated “comprehensively and fairly” as the DOE claims, some issues need to be taken into consideration. The following questions of job satisfaction, working conditions, high qualifications and teacher retention, need to be addressed. “Justice *et al.*’s (n.d.) study found that ‘teachers leaving the profession cite low teacher morale, enhanced by school and district pressure for high student achievement on standardized tests’ as rational (p.384)” (Hill; Barth, 2004). Here are a few past statistics to express the broader problems:

Bowler (2003) conducted a survey to determine the impact of NCLB on rural schools. 75% of secondary and 33 % of elementary teachers said that the ‘Highly Qualified’ designation would impact retention. Principals’ comments on the survey:

Teachers have to go back to school to get qualified.

Richer districts are influencing the super qualified into leaving.

The pool of totally qualified is small.

Impossible task to place qualified teachers in each assigned area” (pp. 4-5) (Hill; Barth, 2004)

### Personal Experience and observations

Before concluding my research paper, I would like to share some of my past experiences and observations in a Hawaiian middle school classroom for the last couple of semesters. I spent the last two semesters observing and helping 6<sup>th</sup> to 8<sup>th</sup> graders ELL students at Jarrett Middle School in Palolo Valley. To better understand the impact of the NCLB law on teachers’ retention, workload and stress, I feel I should share some statistics in relation to Jarrett Middle School. According to the department of education, Jarrett Middle School had a total of 237 students enrolled for the 2009/2010 school year, 25 fulltime teachers and a ratio of 9.4 students per teacher. Though it is necessary to understand about student/teacher ratio that according to the DoE:

Student/teacher ratio is calculated by dividing the total number of students by the total number of full-time equivalent teachers. Please note that a smaller student/teacher ratio does not necessarily translate to smaller class size. In some instances, schools hire teachers part time, and some teachers are hired for specialized instruction with very small class sizes. These and other factors contribute to the student/teacher ratio. Based on the “School Status and Improvement Report” as recently as 2007 there were about 280 students enrolled, of which 45 were in Special Education Programs. The ethnicities, according to that same report, for the year 2006/2007 for the different students was composed of Chinese (5.5%), Filipino (5.5%), Part Hawaiian (14.3%), Japanese (15.8%), Samoan (7.4%), Indo Chinese (12.9%), other (30.5%) and White (2.9%). These

numbers relate to the composition of the classes I tutored and observed. The classrooms that I participated in had a ratio of 10 students for 1,5 teacher (1 teacher who was part time also helped) in 7<sup>th</sup> grade for the Language arts, and 4 students in 6<sup>th</sup> grade Language arts for the same teacher. The teacher/student ratio at Jarrett Middle School has varied over the past 20 years, from 13 down to 9.4 with a peak at 14.6 in 1999 (source National Center for Education Statistics, U.S. Dept of Education). It is important that the classroom and school have an environment conducive to student and teacher engagement. The ratio student/teacher is part of that environment. Benchmarks, student/teacher ratios, classroom compositions, school settings and organisation, type of education system, poor socio-economic backgrounds, all encourage or diminish teacher engagement. Finally, the Jarrett Middle School's homepage expresses that: "We ask for your support this school year by keeping track of your child's homework. Students should be using their planner daily. Parents may use the planner to communicate with teachers. Encourage your child to read at least 30 minutes daily."

Which leads me to another dominating factor in student assessments: "the one-sidedness of accountability; students and parents lack accountability (Milloy et al., 2004). Poor attendance, unwillingness to do homework, and poor turnout at parent teacher conferences make a teacher's job harder" (Hill; Barth, 2004). A frightening possibility is evoked by Deborah M. Hill and Marlene Barth with the following statement: "Student scores are but one factor in measuring teacher effectiveness" (2004). Furthermore "the problem with student test scores is that children are not evenly distributed. If you're not careful, you end up with incentives for good teachers to avoid kids in need" (Darling-Hammond, 1998, p.11) (Hill; Barth, 2004). Survival of the fittest? Is that the direction we are headed? Today there is an urgency. Everyone needs to be involved.

Competition is fierce. Not everyone will win, not everyone will be reaching the top, but we should try to organize the best conditions possible in order to give a chance for everyone to succeed. The danger is about rigging placement tests to only keep the best of the best in order to keep securing funding from the government. It creates an elitist game, which is exactly the opposite of its intended goal. Money is again at the forefront of a pressing issue.

“The Race to the Top program is a \$ 4.35 billion initiative funded by the Recovery and Reinvestment Act. Additionally the USDOE is

Budgeting \$2.5 billion to align state curricula with the NGA/CCSSO standards. An additional \$400 million is budgeted for developing related standardized tests and measures (assessments) resulting in an aligned package of standards, curriculum and assessments...in applying the standards, the administration would mandate specific ‘turn around’ strategies for schools that failed to produce what it considers to be adequate standard-based results. These strategies include firing the principal, firing some of the staff, and converting the school to charter school or closing the school(s).” (Mathis, 2010, pp.1-2) (Deville; Deville-Chalhoub, 2011)

Another proof that money is controls decisions over which assessments to use is given by Butler and Stevens as they write that “the trend now in many states, such as California, is to use statewide assessment results as a part of reward-punishment system, rewarding schools that have high test scores with additional funds” (Helfand, 2000) (2001).

### Conclusion and thoughts

I briefly covered some of the main issues surrounding the NCLB laws and its successor, The Race to the Top implemented by the current administration. There is much to say, much to explain as well as to argue about the NCLB washbacks. One washback, according to me, is that “NCLB compelled schools to attend to the instructional needs of ELLs because the legislation held schools responsible for the performance of these students on standardized state-wide assessments (i.e. on both assessments of English language proficiency and on subject matter tests)” (Deville; Deville-Chalhoub, 2011). It leads to distortion of teaching methods. Throughout

this past semester, in addition to learning about language testing, I attended a class about Second Language Acquisition hypothesis and theories. It is a sobering thought to realize how wide the gap between the political, philosophical, ideological world and the realities on the ground is. Frances A. Butler and Robin Stevens detail that same gap by stating that “many politicians and the public at large do not fully understand the impact of this validity issue; yet test results are being used for high-stakes decisions about school programs despite their potential inappropriateness” (2001). How do they view the validity question?

“For students whose proficiency is still developing, the tests often pose significant reading challenges, that interfere with the assessment of the content they have learned, making their test scores invalid as indicators of content knowledge or achievement” (Butler; Stevens, 2001).

Are we headed towards an education on steroids type situation? Could we see tests being rigged? Could magic potions ( as some athletes were caught taking) be distributed to enhance test performances? Will education become a bloodsport in this Race to the Top? Has it already become one? Each paper I read dealt with the NCLB issue from a different angle: the impact on the operational day to day teaching on the ground, the historical and developmental perspective of standardized assessment tests, impacts on the teachers’ engagement, and last but not least, the accountability question. Yet...Yet...I still believe that at its core, the NCLB does mean well despite its numerous flawed points. The goals set are high. They seem impossible to reach, create stress, depression, teachers quitting, students falling behind and parents left unengaged. One story, detailed by Hill and Barth, summarizes the most negative impact state wide assessment might truly have on everyone from the students to the teachers and the parents. It represents the seemingly “Sisyphian” task of educating children on the part of each of the people involved in a child’s education.

“One example reported by Boaler (2003) concerns a California high school, serving low income students, famous for its outstanding, professional and dedicated mathematics department. On independently designed assessments, student consistently outperform students in wealthier schools. The students previously saw themselves as failures, but started to believe they could learn and be smart like the ‘rich schools’ students. Their self confidence was dashed and esteem diminished however, when they were labeled ‘below average’ and ‘underachieving’ on a state mandated SAT-9 test. The math department protested the non-reflective test failed to measure students learning acquired skills by using unfamiliar contexts and language barriers. The result: lower student/parent confidence and demoralized teachers” (2004).

Frustration and anger result from lack of encouragements and the illusion of a lack of progress despite the proof to the contrary. Another interesting fact describes that “over 24 % of core courses (math, science, social studies, English) in middle/high schools are taught by teachers lacking a college minor in those subjects” (Hill; Barth, 2004). These situations do reiterate the question of validity of state wide assessment testing, and its influence on teachers’ engagement. The why behind the what is crucial to the motivation created in the student by an engaged teacher knowing his subject matter, the reasons for teaching it and communicating those reasons clearly to his classroom. Finally, an example of a major factor influencing student achievement is given by Mr R. Barker Bausell from the New York Times:

[from] one famous study, by the psychologists Todd Risley and Betty Hart, found that children of professional families had 8 million more words directed at them per year from ages 1 to 3 ½ than children from poorer families. (April 30<sup>th</sup>, 2011)

Which brings me back to the socio economic background of the families living in Palolo Valley, or other poor neighborhoods throughout the United States of America. In my experience at Jarrett Middle School, those families represent the bases of the student attending those classes there. It is frightening thought to realized how rigged the competition is from the start. Which all leads me back to my initial question: who truly is accountable? And, are the testings valid?

In 2004, the then secretary of education Rod Paige headed a publication from the US Department of Education titled “No Child Left Behind: A Toolkit for Teachers”. In 2005, and then again in 2007, an action guide titled “Using Student Engagement to Improve Adolescent Literacy” was published by Learning Points under the main title of “Implementing No Child Left Behind”. The DOE website guides anyone interested through their *Blueprint for Reform* concerning the future of state-wide assessments, how to implement them, their guidelines, etc... Finally, on the website Accountability Resource Center Hawaii, part of the State of Hawaii Department of Education, many detailed reports of Hawaii state assessments can be found along with the Superintendent’s annual report on the state of our education.

Those are just some of the many sources available to start digging deeper in order to understand the educational problem better. I do not claim here to know the solutions. Should there be some kind of accountability put in place? Yes. Can it be improved. Yes. My real concern is that we (government, teachers, parents, politicians,...) are all on the same team fighting the same common enemy: un-education or poor education. But are we? Do we all have the same goals? Something to ponder further.

### Final Thought

I will end with a quote from one of our in classroom slides:

“The most important consideration in designing and developing a language test is the use for which it is intended, so that the most important quality of a test is its usefulness”

(Bachman & Palmer, 1996, p.17)

References:

- Bausell, R. Barker (2011, April 30). A New Measure for Classroom Quality. The New York Times. Retrieved from <http://www.nytimes.com>
- Berry, Ruth A. W. (2006). Teacher Talk During Whole Class Lessons: Engagement Strategies to Support the Verbal Participation of Students with Learning Disabilities. *Learning Disabilities Research & Practice*, 21 (4), 211-232.
- Bryson, C., & Hand L. (2007). The role of engagement in inspiring teaching and learning. *Innovations in Education and Teaching International*, vol. 44, pp.349-362.
- Bunch, M.B. (2011). Testing English language learners under No Child Left Behind. *Language Testing*, vol. 28 (3), pp.323-341.
- Butler, F.A., & Stevens R. (2001). Standardized assessment of the content knowledge of English language learners K-12:current trends and old dilemmas. *Language Testing*, vol. 18 (4), pp.409-427.
- Deville, C., & Chalhoub-Deville, M. (2011). Accountability-assessment under No Child Left Behind: Agenda, practice, and future. *Language Testing*, 28 (3), pp.307-321.
- EdisonLearning, Inc. (2011). Retrieved from [http:// www.edisonlearning.com/about-edisonlearning](http://www.edisonlearning.com/about-edisonlearning). May 9, 2011.
- Hill, Deborah M., & Barth M. (2004). NCLB and teacher retention: who will turn out the lights. *Education and the law*, vol. 16, pp.173-181.
- Kennedy, Cathleen A. (1998). *Turning the Tables: Engaging Teachers in the Learning Process*. Retrieved from <http://bearcenter.berkeley.edu/publications/paper2.html>. May 9 2011.
- Learning Point Associates. (2007). *Using Student Engagement to Improve Adolescent Literacy*. (ED -01-CO-0011). Washington, DC: U.S. Department of Education.
- National Center for Education Statistics, U.S. Department of Education (2006-2011). [Graph illustration the Student/Teacher ratio for William P. Jarrett Middle School]. Retrieved from <http://www.schooldigger.com/go/HI/schools/0003000218/school.aspx?entity=20#aboutEntity>. May 9, 2011.
- Read, K (n.d.) *Online Cultural Differences*. LERN 10. Online Tutor Training Project. Retrieved from [http:// www.ccsf.edu/Services/LAC/Lern10 Online Tutor Training/culture.html](http://www.ccsf.edu/Services/LAC/Lern10 Online Tutor Training/culture.html). May 9 2011.



Read, K (n.d.) Online Learning Styles. LERN 10. Online Tutor Training Project. Retrieved from [http://www.ccsf.edu/Services/LAC/Lern10\\_Online\\_Tutor\\_Training/newlearnsty.html](http://www.ccsf.edu/Services/LAC/Lern10_Online_Tutor_Training/newlearnsty.html). May 9 2011.

RMI Embassy (2005). Retrieved from <http://www.rmiembassyus.org/index.htm>. May 9, 2011.

Schussler, Deborah L. (2009). Beyond Content: How Teachers Manage Classrooms to Facilitate Intellectual Engagement for Disengaged Students. *Theory Into Practice*,48, 114-121.

Student engagement. (n.d.). Retrieved May 9<sup>th</sup>, 2011 from the Student engagement Wiki: [http://en.wikipedia.org/wiki/Student\\_engagement](http://en.wikipedia.org/wiki/Student_engagement)

W.P. Jarrett Middle School (2009). Announcements. Retrieved from <http://jarrett.k12.hi.us/home.html>. May 9 2011.