

A CLOSER LOOK ON THE ADMINISTRATION OF STANDARDIZED TESTS IN THE PUBLIC SCHOOL: BASIS FOR POLICY REVIEW

LARA G. TONGOL-DAVID
Researcher

*Master of Arts in Education
Major in Educational Management*

Abstract

This descriptive study aimed to derive inputs as guide on the review of policies on the administration of quarterly standardized tests in public schools. Specifically, it sought to describe the government mandates relative to the administration of standardized tests and identify the benefits and difficulties of the practice based on the perspectives of the school heads, seasoned public school teachers and supervisors of the Department of Education (DepEd). Five school heads and five seasoned teachers from the public schools were interviewed. A separate review on the policies regarding the administration of quarterly tests was conducted. Findings reveal that the DepEd Memorandum Order 522, series of 2009 and DepEd Memorandum Order 55, series of 2016 served as the government mandates relative to the administration of standardized test. An identified difficulty in administering standardized test is that teachers cater to learners with different kinds of capabilities, thinking and levels of understanding. Among the major weaknesses of standardized testing is its inability to detect other equally essential competencies and skills possessed by the students such as creativity and innovativeness. Standardized tests are found to be too narrowly defined and are seemingly limited within the recall dimension of subject content and thus, may not be sufficient measures of a wide range of human potential abilities. It is concluded that standardized testing may be a means to achieve the pursuit of national interest for high academic achievement among students in the context of teachers' and schools' accountability as far as mass assessment is concerned but, on the hindsight, have several shortcomings in dealing with holistic measurement of learners' ability beyond rote learning and memorization. Other aspects of human ability cannot be measured by standardized testing through pencil and paper method. Those skills that are demanded from the 21st century learners such as critical thinking, creativity and collaborative and cooperative adeptness cannot be measured directly by standardized testing. It is recommended that there may be a thorough and comprehensive approach in addressing the implementation and administration of standardized testing in the public schools.

Keywords: Standardized test, teachers, school heads, government mandates

INTRODUCTION

The 21st-century learning environment has a pressing demand for learners to become globally competitive in terms of academic achievement. Educational legislators view this matter with utmost urgency and continue to establish measures such as intensifying the students' competencies by promulgating policies relative to measuring students' academic performance through standardized testing. Policies for quality education channeled through standardized testing are meant to hurdle exams as early as the elementary level to college admission test and eventually pass a board examination.

Wellman (2001) observed that the trend for the adoption of K-12 based programs among countries has radically changed the educational system starting from the primary up to the tertiary level. Senior high school students are eventually being prepared to pass college entrance examinations administered through standardized testing.

Standardized testing, which has been present for almost two centuries, is an essential indicator of students' intelligence in academic learning concerning content and context (Vey, 2005). Therefore, assessment is viewed as one of the central actors in the educational process and a determinant of schools' performance (William, 2010).

Schools, colleges, and universities struggle to adopt more rigorous levels of standardized exams to measure students' performance (Gallagher, 2003). This is in the midst of a globally competitive society where academic achievement is tied to higher performance in standardized testing. Both schools and learners experience pressure in boosting their performances on the national ranking level (Moon, Brighton & Callahan, 2003). Producing students with exemplary performance in standardized examinations has been a critical indicator of school performance for a long time (Mulford, 2003). For private schools, these tests promote an excellent marketing image towards the public (Macbeath, 2005).

The central role of the educational system, as a whole, is to develop people to reach their full potential and become productive citizens of their respective countries. The role of education is widely seen as one of the most promising paths for individuals to achieve a more decent standard of living, along with the macro intention of the government that it can significantly contribute to national economic development (Glennerster et al., 2011).

However, there have been many dissenting perspectives on the aspect of human capital development coming from different angles. One of the most orthodox measures of human capital competency can be traced back as early as the dawning of the first industrial revolution where people were educated to fit their skills needed for the assembly line production system (De Pleijt, 2016; Galor, 2011). This is an essential ingredient in the growing movement when useful knowledge is the key to economic stimulation (Mokyr, 2002, 2009, 2013). This was further expounded by Franck and Galor (2015) in their study.

Thus, it can be hypothesized that the social and economic development of human society is the development of an educational system intended to address the need to produce effective and efficient manpower. It is essential to produce effective and efficient manpower for the support of social and economic development. The educational legislators thought of standardizing learning outcomes, including the implementation of standardized testing geared towards competency requirements of the past and current social and industrial revolution.

The pressing need to produce better graduates has been so apparent in recent years among schools, both for public and private, as a measure of their teaching efficacy, as well as the overall barometer for the country's competitiveness in the human capital market. Hence, the increased reliance and adherence to standardized testing has been adopted across the country as the sole indicator of learners' competency and school's performance (Baker et al., 2010; Pont et al., 2008).

There are numerous literatures on standardized tests, both internationally and locally. However, a specific study that concentrates on how these standardized tests can serve as bases for a policy review. It is imperative to have comprehensive research that deals with this test conducted in public schools and may be utilized by educational managers and teachers.

Standardized testing has been around the humanity's history since the time of Imperial China, where applicants coming from the ruling class, vying for government positions, are given civil service examinations (Fletcher, 2009). During the Tang Dynasty of 618-907 A.D., most of the Confucian literates were selected based on civil service examinations (Fletcher, 2009). These people with the best talents help contribute to the betterment of public administration during that time.

During the 1800s, Horace Mann presented the concept of using standardized exams in Boston schools deriving objective information on the quality of teaching and learning activities among urban schools (Scott, 2004). This assessment helped develop comparative analysis on the individual performance of schools, such as how they deliver their lessons based on students' knowledge and their abilities to answer the exam (Scott, 2004).

At the turn of the 19th century, schoolhouses in the United States began to administer a content assessment to their students see if they have mastered the lessons given to them by their teachers (U.S. Department of Education, 2008; Gallagher, 2003). Students who flank the exam are retained, given additional instructions, and reinforced through an oral examination or recitation (Giordano, 2005).

Educational theorists Edward Ross and Frank Bobbit had begun to promote their idea of 'social efficiency model' on learning at the beginning of the 20th century, postulating that learning must be carried out efficiently, no time should be left for waste, and learners must only be taught on the content they will need in their life's role (Kliebard, 2004). Thus, all learning materials must be focused on the intended outcomes of the learners' adult life.

Schools of specialization were born until this modern time; they still serve to produce graduates who will occupy different jobs in the capitalist market society.

During the early 20th century, Alfred Binet developed the Stanford-Binet Intelligence Test that will later form part of what is now known as the IQ Test (Becker, 2003). The outbreak of World War 2 had prompted the United States military to use the Army Mental Tests to determine the best position for their recruits, followed the introduction of the IBM-developed automated test score scanning bubbled cards. The well-known Scholastic Aptitude Test of today was initially introduced in 1926, which contains 315 questions spanning from subjects of vocabulary to mathematics.

There have been many opposing contentions about the real intention of standardized testing, especially in public schools. Hill (2006, 2005) confirms that standardization of public education is merely serving the purpose of the capitalist market in order to select the best human capital that can be of use for such effect without considering learners' handicap and regardless of the learners' ability to absorb the contents of the instructions given to them (Laprairie et al., 2010).

Learners who best suit the capitalist society are those who rank high in standardized tests (Hursh, 2007, 2008; Horn, 2003). Furthermore, conditioning students to rank high in standardized exams is a mean to compete in the globally competitive labor and economic market where the best of the best will prevail (Giroux, 2009).

According to Giroux (2009), students are merely intellectual slaves who are fashioned to become subjects of the imperial capitalist world and become subservient to the command and control system to standardize their learning and limit their knowledge through the method of specialization and reductionist approaches. In turn, learners are just commodities needed for the capitalist market's burgeoning needs, and that their intellectual qualities have to be polished before joining the workforce (Mahiri, 2005; Kenway & Bullen, 2001).

In the early 1980s, standardized testing was to solicit knowledge on the content of the lesson among students, thus, putting the burden upon the shoulders of the students to excel in their examination. Eventually, it evolved to include accountability among teachers and the school regarding students' performance in the standardized exams (Linn, 2000).

Having been colonized by the Americans for half a decade, the Philippine education system can be distinctively viewed as American education patterned in almost every aspect. This includes all its educational theories and practices in managing school systems and its policies and standards governing the public school examination (Carreon, 2005). The formal examination administered by the American colonial regime was through the Monroe Survey of 1925 as a mandate to evaluate the status of the educational system (Magno, 2010). One of its key findings is that secondary education did not prepare students for life. Thus, it was recommended that training in agriculture, commerce, and industry are strongly needed.

The need to improve the basic education level has prompted the Philippine government to participate in the global movement of Education for All (EFA), which was launched in 1990. The subsequent development of the first EFA initiative was augmented by the 2015 National Plan of Action or Phil-EFA 2015. The thrust is to build a much stronger systemic reform to make a functional education system of making every Filipino literate upon finishing his/her basic education (Lapus, 2008).

The educational reform's primary task is to initiate and sustain the assessment movement, especially in public schools. It is deemed essential for the development of culture on student assessment. It will then utilize the results of the assessment as inputs for further structural reforms and continuous improvement.

Subsequently, the Department of Education has deployed the National Achievement Test (NAT), a Philippine-made standardized test through its Memorandum Order 522 series of 2009, to determine pupils' and students' achievement level, strengths, and weaknesses in the five curricular areas namely: Science, Mathematics, English, Filipino and HeKaSi at the end of the school year (Benito, 2010).

The recent and perhaps a long overdue reform needed by the Philippine education system is the implementation of the enhanced basic education program otherwise known as the K-12 Act or the Republic Act 10533, signed into law by the former President Benigno Simeon Aquino III on May 15, 2013 (Acosta & Acosta, 2016). This was a major revamp in the Philippines' old educational system, with only ten years of basic education schooling and the last country in Asia to adopt a K-12 education structure while two other countries are in Africa, namely Angola and Djibouti (Sarmiento & Orale, 2016).

It is expected that K-12 graduates will have higher academic productivity because of their more prolonged exposure to instructions and are more competitive being exposed to several standardized testing such as the NAT and be at par with the rest of the graduates around the world (Lauraya et al., 2013).

The choice for cost and effort-effective measure in evaluating students' and pupils' academic achievement is by way of administering standardized tests to a multitude of takers at one time. However, since the passage of the Education for All (EFA) and the No Child Left Behind programs, there have been many variable considerations regarding the effectiveness of standardized testing in the extent of its effect on authentic learning and competency accumulation beyond the recall assessment of subject content (Chudowsky et al., 2007; Price, 2010). As modern society is entering the 21st century, learners are expected to showcase more than just recall knowledge by their prospective and future workplaces. Learner's life-long skills like critical thinking, creativity, and collaboration are much more in demand by the real world than recall knowledge. The Philippine government through the Department of Education, should ponder the need for timely and periodic evaluation regarding the purpose and extent of standardized testing that has served the educational landscape. One way of

doing this evaluation is to conduct a discrepancy analysis framework assessment such as the Provus Model.

The very recent adoption of the K-12 education framework has stirred national interests and policy debates among legislators on crafting viable and sound assessment mechanisms that can capture real students' achievement and the level of quality education offered in public schools. Along with this issue is the evaluation of the furtherance of necessary training that public school teachers must need to deliver quality instruction among students and pupils (Duncan, 2010).

The contention that students' academic performance can be measured objectively using standardized examinations has been accepted since its historical inception. It is positively associated with the level of quality of instruction of the school and of teachers delivering the lesson to students. This was amplified by the claims of Heck (2008), Goe (2007), Aaronson et al. (2007), Cochran-Smith and Zeichner (2005), and Rivkin et al. (2005) agreeing on the postulate that there is a significance of teachers' influence on students' outcomes.

The present study's overarching aim is to offer substantive, evidence-based inputs to the Department of Education as a possible groundwork for its policy enhancement program in the process of administering standardized testing in public schools. The inputs generated from this study may also serve as a basis for policy review mechanisms for the legislative branch of the government in considering the adoption and use of standardized testing, and its further development and enhancement.

Specifically, this study addresses the following objectives: (1) identify government policies that have been implemented and promulgated as bases for the adoption and administration of standardized test; (2) describe the perceptions in the practice of administering standardized testing in the public schools as viewed by school heads and teachers as the key informants of this study; (3) describe their offered additional inputs to policy review which may further enhance the implementation and development of standardized testing in the country, and (4) identify the discrepancies found in the policy-implementation of the Department of Education's deployment of the National Achievement Test (NAT) - a Philippine standardized test through its Memorandum Order 522 series of 2009 using the Provus Discrepancy Analysis Model.

It is hoped that the results of this study may also guide teachers, curriculum developers and school administrators in striving to strengthen initiatives to challenge existing assessment and evaluation policies and practices. May other researchers find this study worthy of serving as a starting point for exploring emerging issues on administering standardized examinations and gather more pieces of evidence that may back up the need to evolve practices relevant to standardized testing in the country.

METHOD

The study utilized the descriptive-qualitative research design as the main research framework to solicit necessary data for the inputs to policy review on the administration of standardized tests in basic education sector. Descriptive research provides a precise account of features on a specific individual, state or group; these studies are a way of discovering new insights and describing what happens and/ or categorizing information (Dulock, 1993). Usually, descriptive research method includes questioners, survey, and interview (Koh & Owen, 2000).

Axial coding of key informants' responses was generated to streamline the answers and derive the underlying contexts out of their varied responses. The main process used in the identification of themes is to carefully read the textual responses of key informants and their interrelationships.

The researcher's ability to investigate the variability and associability of the responses reinforced by readings of related literature is the central aim to arrive at the commonality of thematic responses.

Five school administrators and five schoolteachers were selected as respondents of the study and participated in the focus group discussion. The key informants were purposively selected based on their willingness, availability and the specific inclusion criteria set anchored on the objectives of the study. The profile of the key informants is presented below in Table 1. Both school heads and teachers were selected based on their years of experience as educators and on their educational attainment. The school heads covered have at least 20 years of working experience and have a master's degree. The teachers, on the other hand, should have reached the graduate level of education of at least 18 units and must have been an educator for the last 15 years.

Table 1
Participants' profile

School Heads	Years of Experience	Educational Attainment
SH 1	25	M.A. with Ph.D. units
SH 2	22	M.A.
SH 3	21	M.A.
SH 4	20	M.A. with Ph.D. units
SH 5	20	M.A.
Teachers		
T1	16	with M.A. units
T2	20	M.A.
T3	18	with M.A. units
T4	15	with M.A. units
T5	18	with M.A. units

An interview guide was utilized to gather necessary inputs for the qualitative description of the administration of standardized test using a focused group discussion approach. Seasoned public-school administrators were consulted for the content validation of the interview guide so as to confirm that the questions can really solicit the desired inputs for the study.

These are the guide questions utilized to solicit vital information from key informants: (1) What do you think are the main reasons why students must be subjected to standardized assessments? (2) Do you think that standardized testing can benefit both teachers and students? In what way? (3) What are the major drawbacks of teachers and students in administering standardized testing in public schools? (4) Do you think that students' learning is directly measurable by standardized tests? (5) Is there a need for revising, enhancing or replacing standardized testing by any other method of assessment of students' learning?

Narrative responses from key informants were cataloged and were given corresponding coding assignments. School heads were assigned with SH code and teacher participants were given a T code. Hence, the first school head was given an SH1 code and the first teacher informant was given a T1 code. This coding process was applied accordingly for other key informants.

Responses from key informants were carefully noted and processed through axial coding to generate themes for the study. Both deductive and inductive investigation process were used to derive the contexts of the responses.

Confidentiality of key informants' personal information was withheld from the public. Prior to the conduct of the focus group discussion, permission was sought by means of written communication from the school administrators with regard to the intention of the investigation.

Both open and axial coding method were used on the identification of themes. Open coding is concerned with identifying, naming, categorizing and describing the phenomena found in the textual responses. Basically, lines, sentences and paragraphs were read and analyzed in search of the answer to the repeated question about the phenomenon. In this case, the phenomenon is the administration of standardized exams in the basic education sector. Axial coding is the process of relating codes such as categories and themes to each other, through a combination of inductive and deductive analysis. This process is utilized to emphasize causal relationships and fit things into a basic frame of generic relationships as a means of simplifying the variety of textual responses coming from key informants.

To address the fourth objective, this research also used the Provus Discrepancy Analysis Model to evaluate the policy-implementation of the Department of Education's deployment of the National Achievement Test (NAT) - a Philippine standardized test through its Memorandum Order 522 series of 2009. The discrepancy evaluation model, which was developed by Malcolm Provus in 1972, is a useful tool in understanding discrepancy between policy deployment and implementation particularly in a complex organizational structure such

as the Department of Education. A program such as the deployment of standardized testing is examined through its developmental stages which Provus (1972) defines as design, installation, process, and product and cost-benefit analysis. In this manner, the evaluation of such program can be analyzed per stage whether each stage is working or not. The model helps policy makers to draw decisions based on the difference between preset standards and what exists or happens in the implementation of the program.

The evaluation framework in Provus considers discrepancies to be essential clues in educational program evaluation as inputs for further improvement. The Provus' Discrepancy Evaluation Model (PDEM) goes through the four-stage process of analyzing the gap between the intended outcome of standardized testing against the expected competencies that learners gained when they have undergone on such process. The research was anchored on the qualitative analysis of standardized testing from Memorandum Order 522 series of 2009 of the Department of Education which aims to determine pupils' and students' achievement level, strengths and weaknesses in the five curricular areas namely: Science, Mathematics, English, Filipino and HeKaSi at the end of the school year. This was the first stage of the PDEM and serves as the overarching goal of the study. The second stage is the conduct of focused group discussion among selected school heads based on pre-determined selection criteria stated in this Method section. The third stage of the PDEM is to develop themes out of the key informants' responses. The themes were derived from the grouping of similar concepts supplied by the key informants that were subjected through open coding, closed coding and axial coding analysis. The fourth and the last stage is to unravel the discrepancies between the implementation of standardized testing and the actual outcome as observed by the key informants who are experts in the delivery of assessment process, having been in their profession for a considerable number of years and had been seasoned by their work experiences.

FINDINGS

Policies Implemented and Promulgated for the Adoption and Administration of Standardized Test

DepEd Memorandum Order 522 series of 2009. The Department of Education has deployed the National Achievement Test (NAT) - a Philippine standardized test through its Memorandum Order 522 series of 2009, to determine pupils' and students' achievement level, strengths and weaknesses, in the five curricular areas, namely Science, Mathematics, English, Filipino and HeKaSi (Heograpiya, Kasaysayan at Sibika) in elementary and Araling Panlipunan in secondary level at the end of the school year (Benito, 2010). The National Achievement Test was designed pursuant to DepEd Order No. 15 series of 2006, as stipulated under Every Child a Reader Program (ECARP). The NAT, which is primarily a system-based assessment intended to measure learning outcomes across target levels in identified periods of basic education, was one of the locally contextualized and normalized tests in the Philippines.

In particular, it covers the period from mid-assessment of elementary education, which falls on the third grade to the terminal exit assessment at the sixth grade. Results in Grade 6 will serve as a yardstick of incoming first year students' readiness for their high school life. The second-year high school assessment serves as the mid-level assessment for the students' secondary schooling proficiency.

DepEd Memorandum Order 55 series of 2016. Through the DepEd Order 55 series of 2016, 'Policy Guidelines on the National Assessment of Student Learning for the K to 12 Basic Education Program'. This latest policy announcement is geared toward measuring learners' progress in attaining learning standards and 21st-century skills. The result of such adopted policy on standardized testing will form part of the enhancement process of teaching and learning as the country is moving toward globalizing its educational system.

The policy is also intended for accountability measures among schools to gauge their effectiveness and efficiency in delivering an outcomes-based education system that was recently adopted due to the K-12 program. Furthermore, it is a mechanism to determine whether learners are meeting learning standards of the curriculum and measure students' aptitude and occupational interest for career guidance and career choice.

Perceived Benefits in the Practice of Administering Standardized Testing in Public Schools as Viewed by School Heads and Teachers

Objective measurement. The main goal of standardized assessment is to measure students' knowledge based on the academic domain such as language, science, and mathematics. Those administering the standardized test assumed that the test's content could be answered by all takers, considering that students were previously given instructions regarding the content of the test. With this uniformity, a certain measure of fairness and objectivity is achieved, and it is believed that elements of bias are removed.

Efficient method of evaluation. Tests that are proven to provide valid and reliable inferences are then norm-referenced, which means the student's knowledge and skills can be compared to a national sample of students in the same grade level. The effectiveness and efficiency of standardized tests for evaluating students have been the basis among policymakers in evaluating the school system's effectiveness in general and the teaching-learning process inside the classroom, in particular. This also provides inputs to better school management and the further enhancement of the curriculum.

Perceived Difficulties in the Practice of Administering Standardized Testing in Public Schools as Viewed by School Heads and Teachers

The changing educational landscape also calls for the reconsideration of traditional standardized testing. Assessment is continuously changing as it is aligned with the changes implemented in both teaching and learning. Traditional assessment models that include

written assignments and time-constrained examinations merely test students' factual recall of curriculum content. It is now becoming more apparent that assessment must be a fit-for-purpose enabling measurement of the extent to which students have learned together with measuring the extent to which they can demonstrate their learning. Thus, newer modes of assessment focus on evidence of student achievements such as portfolios, role-plays, case-based studies, oral and poster presentations rather than on mere regurgitation of memorized facts and figures.

Assessment does not fit all students. The tests do not perfectly match to the individual needs of the students.

A direct measure of students' learning. Standardized tests are commonly used to measure students' cognitive ability to review their achievement concerning a common body of knowledge associated with a discipline or program. Tests are traditionally used in assessment programming to measure whether students have acquired a specific process- and content-related knowledge.

Standardized testing has its own limitations. Scores on most standardized achievement tests are only limited measures of the latent construct of interest, some aspects of student proficiency. As measures of these constructs, test scores are generally incomplete. They are fallible in two senses: in the traditional statistical sense that they include measurement error, and in the sense that they are vulnerable to corruption or inflation.

Additional Inputs to Policy Review

Participants offered additional inputs to policy review which may further enhance the implementation and development of standardized testing in the country.

Adoption of multi-assessment modalities to gage multiple intelligences. The 21st century learning environment requires learners to become adoptive to changes and develop a sense of trans-disciplinary and multi-disciplinary thinking in contrast to the preceding learning paradigms of the 19th and 20th centuries. Thus, with the consideration of multiple intelligences it is important that gaging learners' knowledge with the use of only one standardized test may not be enough. It is vital for the cultivation of these varying skills and talents that the Dep Ed policy makers should consider developing and implementing different approaches to capture and gage these 21st century educational paradigm. Even teacher made test as long that it has been well developed and validated by competent school authorities and experts in the field of testing and evaluation can be used to measure different learning styles and thinking styles of learners.

Also, teachers opined that the use of standardized test is not applicable to all learners across varying learning spectrums and outcomes. As such, it diminishes the learners' innate capacity to showcase his or her talents by failing the standardized exam which is not

appropriate for his or her learning styles and thinking styles. Hence, it supports the principals' point of views that there is a need to develop other means of testing aside from the standardized exam.

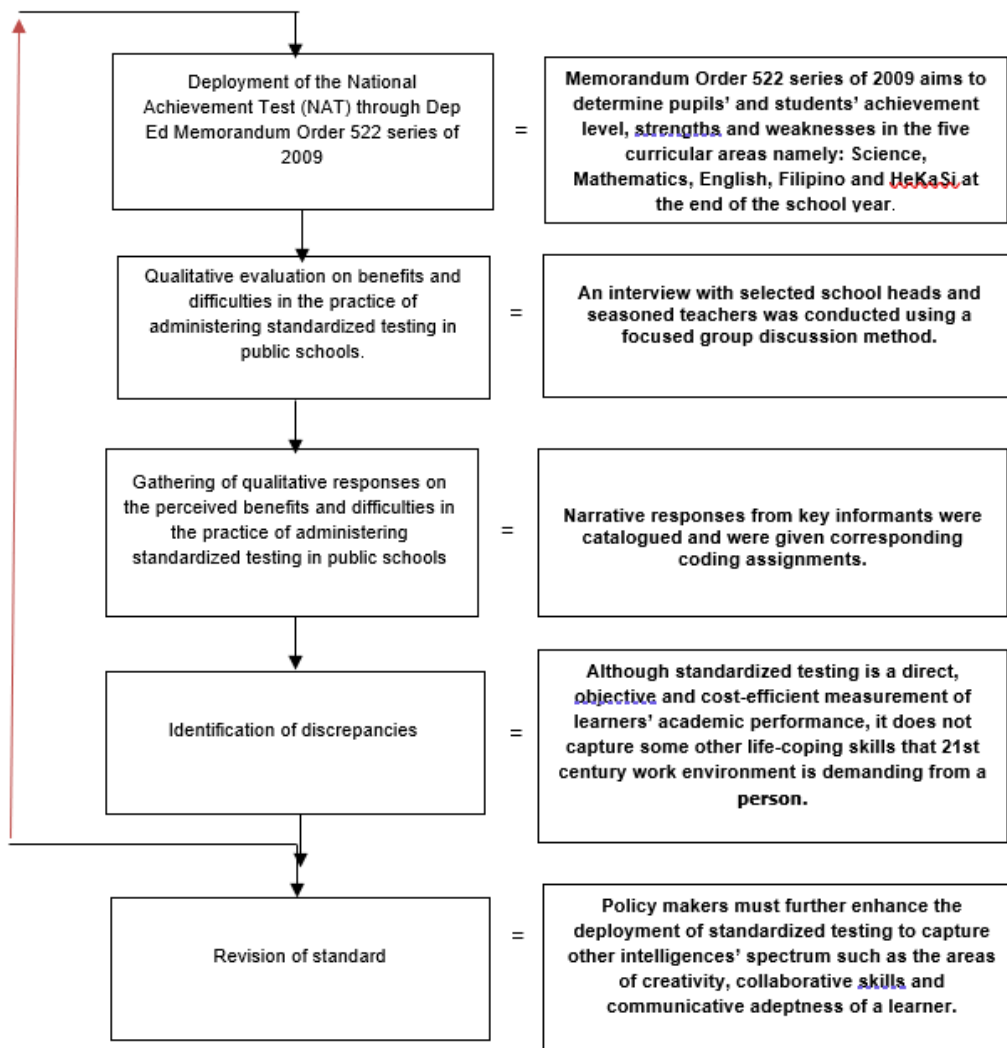
Policy-Implementation Discrepancy Evaluation

Based on the Provus (1972) evaluation model, there are still areas for improvement about the implementation of standardized testing such as the real measurement of life coping skills for the 21st century environment. It was found out that standardized testing can directly and objectively measure pupils' learning out of their academic activities but only to the extent of recall knowledge. Also, standardized testing is a cost effective and efficient way of assessing pupils' academic performance. However, it was revealed that the standardized testing approach cannot assess other practical and life-long learning skills particularly that of the 21st century skills which calls for a more pragmatic and innovative assessment method apart from the traditional pencil-and-paper testing process. Adding to the last stage of the PDEM is the recommendation that policy makers must further enhance the deployment of standardized testing to capture other intelligences' spectrum such as the areas of creativity, collaborative skills and communicative adeptness of a learner.

Figure 1 summarizes the evaluation results following the Provus Evaluation Model.

Figure 1

Provus (1972) Evaluation Model



DISCUSSION

Based on the findings of the study, the implementation of standardized testing determines whether learners are meeting the learning standards of the curriculum and to measure students' aptitude based on the academic domain such as Language, Science, and Mathematics. Further, standardized testing is also an objective means to determine pupils' and students' occupational interests of pupils and students in the future. Such proper career guidance program can be crafted that will be suited for the pupils' career choice.

The implementation of the Philippine standardized test through Memorandum Order 522 series of 2009, with the aim of determining pupils' and students' achievement level, strengths and weaknesses, in the five curricular areas, namely Science, Mathematics, English, Filipino, and HeKaSi (Heograpiya, Kasaysayan at Sibika) in elementary and Araling Panlipunan in secondary level at the end of the school year (Benito, 2010) is a commendable step for the Philippine government in its goal of enhancing the level of education in the country particularly in the basic education. On the global and national scale, there has been no strong empirical association between nationwide learning standards and the achievement outcomes on standardized assessment (Tienken, 2008; Troia et al., 2016).

In recent years, the clamor for a more holistic consideration of a person's innate ability beyond the mechanistic cognitive function of memorizing facts has been deemed to be a more comprehensive erudition and not merely an education where people were being transformed into machine-like artifacts in order to fit their skills needed for the production system (De Pleijt, 2016; Galor, 2011).

Thus, it is recommended that educational administrators implement an alternative tool like a teacher-made test to measure the thinking ability and level of understanding holistically.

On the perceived benefits and difficulties in the practice of administering standardized testing in public schools as viewed by school heads and teachers, it was found that standardized assessment is a cost-effective method of gauging students' and pupils' academic performance despite some shortcomings in measuring the totality of a person's innate knowledge or talent.

Standardized testing offers cost and effort-effective measures in gauging students' and pupils' academic performance. The administration of standardized testing in a multitude of takers, especially for public schools where a significant number of pupils and students are present can be considered a viable mechanism. The theoretical underpinning of the social efficiency model can be addressed by administering standardized testing in public schools. However, it possesses a major shortcoming since it cannot be the major basis neither a sole indicator of student academic performance (Behuniak, 2002).

The choice for cost and effort-effective measure in measuring students' and pupils' academic achievement is by way of administering standardized tests to a multitude of takers at one time. For many years, producing students with exemplary performance in standardized examinations has been a key indicator of school performance (Mulford, 2003). For private schools, these tests promote a good marketing image towards the public (Macbeath, 2005).

Since the passage of the Education for All (EFA) and the No Child Left Behind programs, there have been many variable considerations regarding the effectiveness of standardized testing in the extent of its effect on the authentic learning and competency accumulation beyond the recall assessment of subject content (Chudowsky et al., 2007; Price, 2010). There have been several opposing contentions about the real intention of standardized testing, especially in public schools. Hill (2006, 2005) confirms that the standardization of public education is merely serving the purpose of the capitalist market in order to select the best human capital that can be of use for such effect without considering learners' handicap and regardless of the learners' ability to absorb the contents of the instructions given to them (Laprairie et al. 2010). They are merely human commodities needed for the burgeoning need of the capitalist market, and that their intellectual qualities have to be polished before joining the workforce (Mahiri, 2005; Kenway & Bullen, 2001). These issues posed certain limitations on the very nature of standardized testing and to which it is being intended to gauge students' learning.

Among the major weaknesses of standardized testing is its inability to detect other students' competencies and skills, such as creativity and innovativeness. Standardized tests are too narrowly-defined and restricted only for recall dimensions of subject content and cannot simply solicit a wide range of potential human abilities.

Although pupils' and students' academic performance can be gauged objectively using standardized examinations, which have been accepted since its historical inception and are highly associated with the level of quality of instruction of the school, as well as of teachers delivering the lesson to students, caution may need to be taken as well. Empirical research findings have suggested unintended consequences of standardized exams, particularly on the most vulnerable students (Afflerbach, 2005), to fragment the structure of knowledge and have substantially shifted the frame of learning reference from a student-centered to teacher-centered pedagogy (Au, 2007).

For example, in Finland, as one of the top developed countries for education discouraged using standardized exams to drive academic performance among its schools, given the educational policymakers' belief that standardized assessment narrows the curriculum and leads to unhealthy competition (Sahlberg, 2011).

Standardized tests narrow the entire curriculum in many schools, often squeezing-out subjects such as music, art, foreign languages, elementary grades, and social studies because they are not included in tests. Standardized tests also limit the writing ability among

students. Many standardized tests, such as multiple-choice do not include a significant portion of direct measures that require students to write extended prose. Standardized tests only encourage to emphasize an approach that focuses on applying a fixed set of skills, which means that students learn little about processes of composing and rhetorical dimensions such as audiences and purposes for writing. This limitation is exacerbated by the increasing reliance on standardized tests on machine scoring.

The very recent move by the Philippine government to adopt a K-12 education program for its basic education has stirred national interests and policy debates among legislators on the crafting of viable and sound assessment mechanisms that can capture real students' achievement and the level of quality education offered in public schools.

Additional inputs to policy review to further enhance the implementation and development of standardized testing in the country may be considered by policymakers. Although, producing students with exemplary performance in the standardized examinations has been for many years a key indicator of school performance. As Rothman (2014) notes, "*when there is a discrepancy between tests and standards, teachers tend to place a greater emphasis on what is tested.*" Standardized test items are not parallel with typical classroom skills and behaviors. Since questions have to be generalizable to the entire population, most items assess general knowledge and understanding.

When crafted and designed properly alongside effective implementation and administration, standardized tests can offer students, teachers, administrators, and policymakers substantial benefits. Teachers can take a cue out of the result of standardized test scores, which can help identify areas of strength and weakness within the curriculum. Pupils' and students' academic performance can be measured objectively using standardized examinations and are highly associated with the level of quality of instruction of the school, as well as of teachers delivering the lesson to students. Therefore, standardized assessment is being viewed as one of the central factors in the educational process (William, 2010) and a determinant of schools' performance.

The content not mastered by students can provide teachers and school administrators with valuable information to assess their teaching-learning delivery level (Volante, 2005). Gaps in student learning can give policymakers, school heads, and other concerned stakeholders such as international development groups to give teachers some specific types of professional development to be pursued as part of the continuous improvement process, significantly to strengthen the newly implemented K-12 program.

There shall be a thorough and comprehensive approach framework in addressing the implementation and administration of standardized testing in public schools. At the outset, it can be inferred that the result of the standardized test can reflect the level of teaching and learning process in public schools and the level of instructional delivery and competence among teachers. Aside from standardized testing that is too narrowly defined based on a

predetermined curricular content, there should be other means of alternative assessment that can capture other dimensions of human intelligence. With the consideration of multiple intelligence, it can be generalized that standardization is not a catch-all phrase to address the development of human potential but instead through a mixed and dynamic formulation of the teaching and learning process adaptive to the rapidly changing social, economic, and technological changes.

Due to the changing educational landscape, there must be a reconsideration of traditional standardized testing. Traditional models of assessment that include written assignments and time-constrained examinations merely test students on factual recall of curriculum content. Newer modes of assessment focused on evidence of student achievements such as portfolios, role-plays, case-based studies, oral and poster presentations rather than merely regurgitation of memorized facts and figures are being considered to augment the traditional paper and pencil assessment. Also, all learning materials must be focused on the intended outcomes of the learners' adult life.

There is still a long way to realize comprehensive development in the Philippines' basic education assessment system. The implemented standardized test is only the beginning since it only captures the pupils' lower-level thinking skills, especially the recall domain. The other opportunities for improving and enhancing the assessment process must also capture the other domains of learning, such as the practical assessment method, which can elicit complex demonstrations of learning and measure the full range of knowledge and skills necessary for young learners to succeed in college and 21st-century careers (PARCC, 2010). This is necessary for gauging learners' higher order thinking skills, which cannot simply be captured using a simple paper and pencil assessment system.

Standardized testing is a means to achieve the pursuit of national interest for high academic achievement among students in the context of teachers' and schools' accountability to deliver the curriculum's necessary content with the condition that it must be well-understood among students as essential in gaining high scores during examinations. It is concluded that standardized testing may be a means to achieve the pursuit of national interest for high academic achievement among students in the context of teachers' and schools' accountability as far as efficiency for mass assessment is concerned but, on the hindsight, have several shortcomings in dealing with holistic measurement of learners' ability beyond recall knowledge, rote learning, and memorization. Other aspects of human ability cannot be measured by standardized testing through the pencil and paper method. Those skills demanded from the 21st-century learners, such as critical thinking, creativity, and collaborative and cooperative adeptness, cannot be measured directly by standardized testing.

Policy makers on this regard may consider looking beyond the mere measurement of rote knowledge recall that is being captured by the standardized paper and pencil test mechanisms. The aspects of feedback mechanisms to learners as well as to the teachers must be developed from the top hierarchy of Dep Ed to provide appropriate enhancement

measures on how to improve the performance of learners and teaching strategies of teachers on the next testing period and eventually leading to improved test scores performance. Further, it has been well noted that standardized testing is still far from developing learners' competencies to becoming competitive in the 21st century environment along with the acquisition of necessary skills such as creativity. This issue may be considered and reevaluated deeper by policy makers on how capture and develop appropriate assessment packages that can measure such realities especially that the world has entered a knowledge society. The recent international assessment results comparing the competency index of learners among participating countries especially in the areas of math, science and English language shows that the Philippines is lagging behind in these three areas (PISA, 2019; TIMSS, 2019). This may be a point of realization and reflection among policy makers that the Philippine educational system is not yet at par with the developed countries. As such, the study that the testing mechanisms of the Philippine educational system is way beyond other countries' testing mechanisms, judging from the results of these international comparison indices.

Also, policy makers may consider studying the appropriateness of standardized testing in the context of learners' diversity. Considering that learners come from diverse backgrounds in terms of culture, tradition, English language proficiency, socio-economic status and even heterogeneity in learning and thinking styles. Standardized testing on that regard should not only be developed as a single normative reference but must also be considered in the aspects of group norms. Furthermore, the content of standardized testing shall be aligned with the demands and requirement of the greater society and industry to appropriately prepare learners for real life knowledge instead of book knowledge which are sometimes detached from the reality. Aside from the paper and pencil test, policy makers can further extend the development of a more realistic assessment such as industry-based aptitude exam as well as actual of practical assessment modalities. The content of the standardized test must be based on real life context of social issues and the dynamism of social, political, technological, and economic realities of the modern world. The main principle of education is not simply to pass the standardized test but rather to prepare the learners to become competitive, resilient, and productive citizens of a country. Thus, the aim of both policies focused on simply making learners pass the exam must be changed and should be replaced by the paradigm of a more responsive and innovative methods of gaging learners' real knowledge that can be used productively in this dynamic and ever changing 21st century environment.

The inference drawn from this study is hoped to initiate the rethinking of the existing policies, procedures, implementing guidelines, and practices in the deployment of standardized testing to reexamine the content and context of the program. For the conclusions of the study, standardized testing was revealed to: (1) promote teacher-centered pedagogy; (2) determine students' occupational interests; (3) objectively measure students' aptitude by assessing general knowledge and understanding; and, (4) to cover a limited range of potential

human abilities that mostly focus on lower-level thinking skills and miss to gauge the essential 21st century skills such as creativity, and the abilities to collaborate and innovate.

After the thorough analysis of the findings of the study, this study underscores the following recommendations: (1) There shall be a thorough and comprehensive approach framework in addressing the implementation and administration of standardized testing in public schools. At the outset, it can be inferred that the result of the standardized test can reflect the level of teaching and learning process in public schools and the level of instructional delivery and competence among teachers. (2) The gaps in student learning as reflected in the results of the standardized test scores can give policymakers, school heads, and other concerned stakeholders such as international development groups to give teachers some specific types of professional development to be pursued as part of the continuous improvement process, significantly to strengthen the newly implemented K-12 program. (3) There should be other means of alternative assessment that can capture other human intelligence dimensions, such as creativity and innovativeness. (4) The standardized tests may be re-calibrated in such a way that these would be student-centered and not teacher-centered. Newer modes of assessment focused on evidence of student achievements such as portfolios, role-plays, case-based studies, oral and poster presentations may also be adopted. (5) Standardized testing can help in determining the students' occupational interests. Given this, the guidance program should be strengthened to guide the students in their career choice, as reflected in the standardized test results. (6) Standardized test questions may be focused on actual classroom skills and behaviors instead of general knowledge. (7) Teachers, school heads, and policymakers may take standardized test scores to help them identify areas of strength and weaknesses within the curriculum.

REFERENCES

- Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics* 25 (1), 95-135. <https://doi.org/10.1086/508733>.
- Acosta, I. C. & Acosta, A.S. (2016). Teachers' Perceptions on Senior High School Readiness of Higher Education Institutions in the Philippines. *Universal Journal of Educational Research* 4 (10), 2435-2450. 10.13189/ujer.2016.041024.
- Afflerbach, P. (2005). National Reading Conference policy brief: High stakes testing and reading assessment. *Journal of Literacy Research* 37 (2), 151–162. https://doi.org/10.1207/s15548430jlr3702_2.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36 (5), 258–267. <https://doi.org/10.3102/0013189X07306523>.
- Baker, E. L., Barton, P.E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R. J. & Shepard, L.A. (2010). *Problems with the Use of Student Test Scores to Evaluate Teachers*. Economic Policy Institute.
- Becker, K. A. (2003). *History of the Stanford-Binet intelligence scales: Content and psychometrics*. The Riverside Publishing.
- Behuniak, P. (2002). Consumer-referenced testing. *Phi Delta Kappan* 84 (3), 199-207. <https://doi.org/10.1177/003172170208400306>
- Benito, N. (2010). *Achievement Test: An Overview*. Department of Education, National Educational Testing and Research Center. https://studylib.net/doc/8112076/national_achievement_test_dr-benito.
- Carreon, M. L. (2005). *Applicability of standard tests to the Filipino*. University of Michigan.
- Chudowsky, N., Chudowsky, V. & Kober, N. (2007). *Has student achievement increased since no child left behind? Center on Education Policy*, Washington, D.C.
- Cochran-Smith, M. & Zeichner, K. (2005). *Studying teacher education: The report of the AERA panel on research and teacher education*. Mahwah, NJ: Lawrence Erlbaum.
- De Pleijt, A. M. (2016). *The role of human capital in the transition to modern economic growth* [Doctoral dissertation, Purdue University]. Utrecht University, The Netherlands.
- Department of Education. (2016). *Order 55. Policy Guidelines on the National Assessment of Student Learning for the K to 12 Basic Education Program*. <https://www.deped.gov.ph/2016/06/30/do-55-s-2016-policy-guidelines-on-the-national-assessment-of-student-learning-for-the-k-to-12-basic-education-program/>

- Department of Education (2009). *Memorandum No. 522. Administration of the National Achievement Test (NAT) for Grade 3, Grade 6 and 2nd Year High School*. <https://www.deped.gov.ph/2009/12/08/december-8-2009-dm-522-s-2009-administration-of-the-national-achievement-test-nat-for-grade-three-grade-six-and-second-year-high-school/>.
- Dulock, H. L. (1993). Research Design: Descriptive research. *Journal of Pediatric Oncology Nursing* 10 (4), 154-157. <https://doi.org/10.1177/104345429301000406>.
- Duncan, A. (2010, November 16). *Remarks to national council for accreditation of teacher education* [Speech]. National Press Club, Washington, D.C.
- Fletcher, D. (2009). *Standardized testing*. *Time*. Retrieved October 15, 2018, from <http://content.time.com/time/nation/article/0,8599,1947019,00.html>.
- Franck, R. & Galor, O. (2015). *Is Industrialization Conducive to Long-Run Prosperity? Working Papers 2*. Brown University.
- Gallagher, C. J. (2003). Reconciling a Tradition of Testing with a New Learning Paradigm Source. *Educational Psychology Review* 15 (1), 83-99. <https://doi.org/10.1023/A:1021323509290>.
- Galor, O. (2011). *Unified Growth Theory*. Princeton University Press.
- Giordano, G. (2005). *How testing came to dominate American schools: The History of Educational Assessment*. Peter Lang Inc.
- Giroux, H. A. (2009). Obama's dilemma: Postpartisan politics and the crisis of American education. *Harvard Educational Review* 79(2), 250-266. <https://doi.org/10.17763/haer.79.2.x576782u27885513>.
- Giroux, H. A. (2009). The politics of higher education and the militarized academy after 9/11. *Alif: Journal of Comparative Poetics*, 29, 104-126. <https://doi.org/10.1177/0263276408095216>.
- Giroux, H. A. (2009). *Youth in a suspect society*. Palgrave Macmillan.
- Glennerster, R., Kremer, M., Mbiti, I. & Takavarasha, K. (2011). *Access and Quality in the Kenyan Education System: A Review of the Progress, Challenges and Potential Solutions*. Office of the Prime Minister of Kenya.
- Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. National Comprehensive Center for Teacher Quality. <https://eric.ed.gov/?id=ED521219>.
- Heck, R. H. (2008). Teacher effectiveness and student achievement: Investigating a multilevel cross-classified model. *Journal of Educational Administration* 47(2), 227-249. <http://dx.doi.org/10.1108/09578230910941066>.

- Hill, D. (2006). Class, capital and education in this neoliberal and neoconservative period. *Information for Social Change*, (23), 1-19. https://doi.org/10.1057/9780230104709_6.
- Hill, D. (2005). Globalization and its educational discontents: Neo-liberalization and its impact on education workers' rights, pay and conditions. *International Studies in Sociology of Education* 15 (3), 257-288. <https://doi.org/10.1080/09620210500200143>.
- Horn, C. (2003). High-stakes testing and students: Stopping or perpetuating a cycle of failure? *Theory Into Practice*, 42(1), 30-41. https://doi.org/10.1207/s15430421tip4201_5.
- Hursh, D. (2007). Exacerbating inequality: The failed promise of the *No Child Left Behind Act*. *Race, Ethnicity and Education*, 10 (3), 295-308. <https://doi.org/10.1080/13613320701503264>.
- Hursh, D. (2008). *High-stakes testing and the decline of teaching and learning: The real crisis in education*. Rowman & Littlefield Publishers.
- Kenway, J. & Bullen, E. (2001). *Consuming Children. Education-entertainment-advertising*. Open University Press.
- Koh E.T. & Owen, W.L. (2000) *Descriptive Research and Qualitative Research*. Introduction to Nutrition and Health Research.
- Kliebard, H. M. (2004). *The struggle for the American curriculum 1893-1958*. Routledge Falmer.
- Lapus, J.A. (2008). The Education System Facing the Challenges of the 21st Century. Country Report of the Republic of the Philippines, Geneva, Switzerland.
- Laprairie, K., Johnson, D., Rice, M., Adams, P., & Higgins, B. (2010). The top ten things new high school teachers need to know about servicing students with special needs. *American Secondary Education* 38 (2), 23–31. <https://doi.org/10.1177/0731948712451261>.
- Lauraya, F.P., Mascariñas, A., Amano, L., Bercasio, R., Cortez, C. Jr. & Torres, E. (2013). Bicol SUCS after K-12 and Beyond: An Ex-ante Analysis of the Impact of K-12 Program to HEIs, *Higher Education Summit 2013*. <http://scholar.google.com/scholar?cluster=4988710357789747193&hl=en&oi=scholar>.
- Linn, R. L. (2000). Assessments and accountability. *ER Online* 29 (2), 4-16. <https://doi.org/10.3102/0013189X029002004>.
- Macbeath, J. (2005). *Schools must speak for themselves: the case for school self-evaluation*. Routledge.
- Magno, C. (2010). A brief history of educational assessment in the Philippines. *Educational Measurement and Evaluation Review* 1, 140-149. <https://www.pemea.org/volume-1-c4s>.

- Mahiri, J. (2005). From 3 R's to 3 C's: Corporate curriculum and culture in public schools. *Social Justice*, 32(3), 72-88. <https://www.jstor.org/stable/29768322>.
- Mokyr, J. (2002). Long-Term Economic Growth and the History of Technology. *Handbook of Economic Growth* 1, 1113-1180. Elsevier.
- Mokyr, J. (2009). *The Enlightened Economy: An Economic History of Britain, 1700- 1850*. Yale University Press.
- Mokyr, J. (2013). Human Capital, Useful Knowledge, and Long-Term Economic Growth. *Economia Politica* 30 (3), 251-271. <https://doi.org/10.1428/75293>.
- Moon, T., Brighton, C. & Callahan, C. (2003). State standardized testing programs: Friend or foe of gifted education? *Roeper Review* 25 (2), 49-60. <https://doi.org/10.1080/02783190309554199>.
- Mulford, B. (2003). School leaders: Changing roles and impact on teacher and school effectiveness. *Education and Training Policy Division: OECD*. <https://www.oecd.org/education/school/2635399.pdf>.
- Partnership for Assessment of Readiness for College and Careers (2010). Application for the race to the top comprehensive assessment systems competition. *PARCC*. <http://www.fldoe.org/parcc/pdf/aprtcasc.pdf>.
- Pont, B., Nusche, D. & Moorman, H. (2008). Improving school leadership. *Education and Training Policy Division: OECD*. <https://www.oecd.org/education/school/40545479.pdf>.
- Price, H. E. (2010). Does no child left behind really capture school quality? Evidence from an Urban School District. *Educational Policy* 24 (5), 779-814. <https://doi.org/10.1177/0895904810376564>.
- Provus, M. (1972). Evaluation as Public Policy. *Curriculum Theory Network* 3 (8-9), 33-44. <https://doi.org/10.2307/1179190>.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica* 73 (2), 417-458. <https://doi.org/10.1111/j.1468-0262.2005.00584.x>.
- Rothman, R. (2014). The Common Core takes hold. *Education Next*, 14(3), 16–22. <http://educationnext.org/the-common-core-takes-hold/>.
- Sahlberg, P. (2011). Lessons from Finland. *American Educator* 35 (2), 32–36. <https://files.eric.ed.gov/fulltext/EJ931215.pdf>.
- Sarmiento, D. H. & Orale, R. (2016). Senior High School Curriculum in the Philippines, USA and Japan. *Journal of Academic Research*, 1(3) 12-23.

<https://scholar.google.com/scholar?oi=bibs&cluster=1704446097669067795&btnl=1&hl=en>.

- Scott, T. (2004). Teaching the ideology of assessment. *Radical Teacher*, 71(4), 30-37. <http://www.jstor.org/stable/20710280>.
- Tienken, C. H. (2008). Rankings of international achievement test performance and economic strength: Correlation or conjecture? *International Journal of Educational Policy*, 3 (4), 1–15. <https://journals.sfu.ca/ijepl/index.php/ijepl/article/view/110/44>.
- Troia, G., Olinghouse, N. G., Wilson, J., Steward, K. A., Mo, Y., Hawkins, L. & Kopke, R. A. (2016). Common Core Writing Standards: A descriptive study of content and alignment with a sample of former state standards. *Reading Horizons* 55 (3), 98–141. <https://www.jstor.org/stable/90024223>.
- U.S. Department of Education (2008). *Great expectations holding ourselves and our schools accountable for results*. <https://eric.ed.gov/?id=ED504194>.
- Vey, L. D. (2005). *Enhancing the relationship between learning and assessment* [Unpublished doctoral dissertation], University of Canberra, Australia.
- Volante, L. (2005). Accountability, student assessment, and the need for a comprehensive approach. *International Electronic Journal for Leadership in Learning*, 9 (6). <http://ijll.synergiesprairies.ca/iejll/index.php/ijll/article/view/406/68>.
- Wellman, J. V. (2001, March/April). Assessing state accountability systems. *Change*, 33(2), 47- 52. <https://doi.org/10.1080/00091380109601787>.
- William, D. (2010). Standardized testing and school accountability. *Educational Psychologist* 45 (2), 107–122. <https://doi.org/10.1080/00461521003703060>.