doi: http://dx.doi.org/10.7203/RASE.11.3.13054

Págs. 421-432

After the Press Release on Mathematics Achievement: The Alignment of Formative Assessments and Summative Standardized Tests for Students from Minoritized Language Backgrounds

Después del comunicado de prensa sobre el logro de las matemáticas: la alineación de las evaluaciones formativas y las pruebas estandarizadas sumativas para los estudiantes de entornos lingüísticos minorizados

Saúl Isaac Maldonado, Rebeca Judith Andrade[1]

Abstract

Media coverage often reports standardized test scores as indicative of global competitiveness and journalists frequently frame the mathematics achievement of U.S. students using a discourse of failure, particularly for students from minoritized backgrounds. Informed by research literature on: (a) formative and summative assessment, (b) the interdependence of language and mathematics, and (c) a sociopolitical stance for researching students from minoritized language backgrounds, this study analyzed the linguistic alignment of mathematics assessments. Results from this study show how alignment between formative assessment processes and summative standardized tests influences performance on high-stakes exams. Findings call into question the incomplete media coverage of students' achievement trends that perpetuate stereotypes of failure and advocate for a national formative assessment program in the United States.

Keywords

Standardized test scores; formative assessment; bilingual students.

Resumen

Los medios de comunicación a menudo presentan puntajes de exámenes estandarizados como indicativos de la competitividad global y las periodistas frequentemente enmarcan los logros matemáticos de los estudiantes de EEUU utilizando un discurso de fracaso, especialmente para los estudiantes de entornos minoritarios. Guiado por la literatura de investigación sobre: (a) la evaluación formativa y sumativa, (b) la interdependencia del lenguaje y las matemáticas, y (c) una perspectiva sociopolítica para investigar a estudiantes de contextos lingüísticos minorizados, este estudio analizó la alineación lingüística de las evaluaciones de matemáticas. Los resultados de este estudio muestran cómo la alineación entre los procesos de evaluación formativa y las pruebas estandarizadas sumativas influye en el rendimiento de las pruebas escolares de alto riesgo. Los hallazgos cuestionan la cobertura incompleta de los medios de comunicación sobre las tendencias de rendimiento de los estudiantes que perpetúan los estereotipos de fracaso y abogan por un programa nacional de evaluación formativa en los EEUU.

Palabras clave

Pruebas estandardizas, asesoramiento, estudiantes bilingües.

Recibido: 21-03-2018
Aceptado: 28-04-2018

1 Saúl Isaac Maldonado, San Diego State University, smaldonado@sdsu.edu; Rebeca Judith Andrade, Glendale Unified School District, randrade@gusd.net.

1. Introduction

The United Nations Treaty, «Convention on the Rights of the Child,» advocates: (a) the taking of appropriate measures to protect children against all forms of discrimination, (b) the promotion of assistance to develop and preserve children's cultural identity and values, as well as (c) the provision of «access to scientific and technical knowledge and modern teaching methods» (UN General Assembly, 1989, Articles 2, 8, 28.3, 29). Most countries ratified the Treaty's declaration to protect children from discrimination, promote children's cultural identity and provide access to appropriate pedagogy in 1990; however, the United States did not become a signatory until 1995[2]. Although every nation can improve their equity-oriented initiatives for students, schools, and systems (OECD, 2012), this study investigates how fairness and inclusion in mathematics education inform equitable assesssment in the United States (Field *et al.*, 2007).

Policies and practices in U.S. schools are often influenced by aggregated standardized test scores as the primary assessment of students' learning. Standardized test scores are summative indicators *of* students' learning and distinctive from formative indicators *for* students' learning that include processes designed to improve educators' instruction and bolster students' development. Providing «information about current status and progress of student achievement and quality of schooling,» standardized test results often serve as the central catalyst for «reforming educational practices» (Miller *et al.*, 2009: 3). While summative standardized test scores and formative assessment processes are interdependent, many media communications report students' standardized test scores in the aggregate without describing students' individual characteristics, such as degree of English-language background, family and community characteristics, such as socio-economic status, or students' access to formative assessment processes. Formative assessment processes provide educators with actionable information regarding students' thinking and autonomy for learning (Cizek, 2010).

In 1965, the Elementary and Secondary Act (ESEA) established national summative assessment requirements to document «disparities in educational opportunities and in student performance» without corresponding policies for a national formative assessment program (4). In 1983, «A Nation At Risk» reported declines in standardized test scores as indicative of «a rising tide of mediocrity that threatens our very future as a Nation» (National Commission on Excellence in Education: 5). Although the authenticity of results from «A Nation at Risk» have been refuted (Carson, *et al.*, 1992), the «report was released in a White House ceremony» and the media coverage invoked a rhetoric of educational crisis that remains commonplace in contemporary times (Mehta, 2015: 21). In 2001, ESEA was re-authorized as the No Child Left Behind Act (NCLB) and summative assessment requirements have become high-stakes measures of accountability. For schools, standardized tests determine economic rewards and sanctions (e.g., programmatic funding, teacher reassignment, school reconstitution). For students, standardized tests determine curricular experiences (e.g., academic track placement, grade promotion, high school graduation) (Heubert & Hauser, 1999). Historically, high-stakes assessment is designed as economically efficient modes for evaluating student learning, identifying underachieving groups, and allocating incentives and penalties. Critics of high-stakes assessment argue that such modes of evaluation do not provide meaningful evidence of student learning, exacerbate achievement disparities, and create temporary compliance (Kohn, 2000).

2 See United Nations, Treaty Series, vol. 1577, p. 3; depositary notifications C.N.147.1993. TREATIES-5 of 15 May 1993 [amendments to article 43 (2)]; and C.N.322.1995. TREATIES-7 of 7 November 1995 [amendment to article 43 (2)].

The National Assessment of Educational Progress (NAEP), also known as «The Nation's Report Card,» is a U.S. assessment program that measures what students know and can do in various subject areas. NAEP assessments remain the same year-to-year, allowing for decreased variables, and serve as a common metric of achievement for students in all states (National Center for Education Statistics [NCES], 2018). The most recent NAEP results show that 60 percent of fourth graders, and 66 percent of eighth graders, did not score at or above the proficient designation in mathematics. NAEP data disaggregated by students' racial-ethnic background communicates a consistent narrative: Asian students have the highest percentages of proficiency, followed by White students, Latinx students, and Black students (NCES, 2018). U.S. students' achievement on standardized tests is also compared internationally. The Trends in International Mathematics and Science Study (TIMSS) assessment program provides ongoing comparative data of students in over sixty nations (Provasnik *et al.*, 2016). The most recent TIMSS data in mathematics ranked U.S. students as fourteenth in fourth grade and tenth in eighth grade, with consistent higher rankings for students from: Singapore, Republic of Korea, China, Japan, and Russia. Mathematics achievement disparities between 15-year old students across 72 nations are also evident in data from the Programme for International Student Assessment (PISA), where U.S. students ranked 39th (OECD, 2016).

After the press releases of TIMSS, PISA, and NAEP data, media coverage often reports standardized test scores as indicative of global competitiveness and journalists often frame the mathematics achievement of U.S. students using a discourse of failure, particularly for students from minoritized backgrounds. Stack (2007) analyzed how media reports and press releases from the U.S. Department of Education interpreted TIMSS and PISA results to promulgate narratives of national security and meritocracy while neglecting to communicate the resource disparities across students' racial-ethnic background, gender, and social class. Other researchers have contended that media portrayals of education in U.S. media are systematic representations of the poor and minoritized as dangerous (Darder & Mirón, 2007; Rodríguez & Brown, 2009). Darling-Hammond (2010) presents how a national commitment to equity, and preparing educators to use formative assessment processes, could ameliorate opportunity gaps that sustain educational achievement disparities in the U.S. This study is informed by research literature on: (a) the interplay between assessment and instruction, (b) the interdependence of language and mathematics, and (c) a sociopolitical stance for researching students from minoritized language backgrounds.

1.1. Content Exposure, Resource Allocation, and Instructional Practices

While situating assessment in the policy landscape is important, interpreting the results of content-area assessments requires an explicit consideration of students' characteristics, such as language and literacy competencies, as well as participation experiences in complementary opportunities to learn (OTL). Assessment results are useful markers of content-area achievement and function as outputs; however, results may not fully disclose student inputs (i.e., characteristics and experiences). A helpful conceptual framework for communicating the interdependent relationship between assessment outputs and student inputs is OTL.

OTL is desribed as the interactive interplay amongst: content taught, resource allocation and teaching practices (Pullin, 2008). As instructional content, OTL accounts for students' exposure to the content evaluated on assessments (e.g., curricular fidelity, teacher content coverage). As resources, OTL is defined as access to teaching and learning supports (e.g., qualified educators, access to technology,

supportive services and expenditures). Finally, as processes, OTL is associated with the values placed on instructional practices by educators. In other words, OTL is a composite of content, resources, and processes that integrates student background inputs and is often ignored when interpreting summative assessment outputs (Pullin & Haertel, 2008).

Beyond measures of technical quality, assessment outcomes must be considered in relation to OTL. OTL offers a complementary conceptual framework for interpreting assessment results. In fact, «Schools exist to offer OTL; the study of OTL is the study of schooling» (Pullin & Haertel, 2008: 34). For instance, OTL accounts for specific student background inputs, such as access to credentialed teachers and participation in advanced coursework as well as indicators of resource allocation. Further, OTL considerations raise important concerns, mainly: do assessments do more than reflect content, do assessments become the content? Additionally, it is also important to highlight that the variability in individuals' OTL should not impede policy and practice reconsiderations, for variability exists across definitions of curricular content and endorsed curricula across states and school districts (Pullin & Haertel, 2008).

1.2. The Language Load of Mathematics

Learning the language of the mathematics discipline requires the interdependent development of both content and process dimensions (Bullock, 1994). Shwartz & Kenney (1995) qualify numbers, measurements, shapes, spaces, functions, patterns, data and arrangements as content strands, or mathematical nouns or objects. In terms of mathematical verbs, Shwartz & Kenney (1995) describe modeling and formulating, transforming and manipulating, inferring and communicating, as the predominant problem-solving and reasoning mathematical actions. As students learn the language of mathematics, they must use appropriate mathematical nouns and verbs, not solely read and write terms, definitions, procedures and properties.

Reading in mathematics requires familiarity with conceptually dense prose, decoding of both numeric and non-numeric symbols, and graphical interpretation as well as the ability to categorize distinctions between supportive prose, explanatory information and problem statements (Metsisto, 2005). To support students' learning, mathematics teachers may modify reading strategies from other disciplines, such as strategic reading (Draper, 2002), the Frayer Model (Frayer, Frederick, Klausmeier, 1969), the Semantic Feature Analysis Grid (Baldwin *et al.*, 1981), and guided reading (Allen, 2003). Similarly, writing in mathematics is discipline-specific and embedded in logical-mathematical intelligence, distinctive from linguistic intelligence (Gardner & Hatch, 1989). In mathematics, writing communicates conceptual understanding of relationships rooted in the material world and provides an instructional opportunity for developing and evaluating computational development (Tuttle, 2005). While writing fictional and historical narratives of mathematicians and math concepts would complement mathematics teaching and learning, it would be more appropriate to develop students' proficiency in appropriate expository texts that justify strategies and explain answers, not solely solve an equation. To support students' learning, mathematics teachers may modify writing strategies from other disciplines, such as response logs (Maloch, 2002), word recordings (Lappan *et al.*, 2002), and structured guides to structure writing processes that document mathematical thinking (Fisher, Frey, & Williams, 2002).

Learning environments where mathematics teachers provide instruction with authentic literacy opportunities procure dual benefits, students «communicate to learn mathematics, and [students] learn to communicate mathematically» (National Council of Teachers of Mathematics, 2000: 60). In a learning environment where mathematics communication is authentic, students and teachers use key math terms (e.g., add, length, place value, fraction, area, quotient, dividend, coordinate, variable) as well as process terms (e.g., measure, estimate, solve, convert, graph, evaluate, represent). Further, teachers provide substantive instructional attention to students' development of both key and process terms (e.g., student glossaries, word walls, structured writing tasks, reading clarifications), check for student accuracy in use of terminology and promote students' use of key words in productive (speaking, writing) language.

The usage of consistent mathematics terminology between formative assessment processes and summative standardized tests is necessary but insufficient for creating equitable contexts for mathematics learning. Equally important to consistent terminology on assessments are the instructional literacy practices teachers use to develop students' mathematical proficiency as well as the resources students are provided to engage in literacy tasks. Authentic literacy tasks in mathematics include: identification, description and comparison of measureable attributes, interpretation and representation of data, and generation and analyses of patterns. Appropriate resources to engage in authentic mathematical literacy practices include objects, shapes, figures, and graphs. In a learning environment where authentic literacy practice is evident, teachers address numerous mathematical practices, including abstract quantitative reasoning, argumentation, communicative appropriateness, and resource selection (NGACBP & CCSSO, 2010: 6-7).

1.3. Equitable Mathematics Education for LM Students

Demography shows that a significant population of U.S. schools are comprised of bilingual students minoritized by their language background (LMs). In 2016, 22 percent of children and youth ages 5 to 17 in the United States (or 12.1 million children and youth) spoke a language other than English at home (U.S. Census Bureau, 2016). In the past, this student population was concentrated in states such as California, Texas, Illinois, Florida and New York, but significant trends in national demographic shifts of LMs have been identified using four comprehensive federal datasets, including county-specific data from the Local Education Agency Universe Survey. From 2000 to 2010, states that experienced an increase exceeding 200 percent of LMs included: Mississippi, Kentucky, Indiana, Virginia, South Carolina and Delaware (National Clearinghouse of English Language Acquisition, 2011).

Gutiérrez (2007) describes equity in mathematics education as an «*inability to predict* student outcomes based *solely* upon cultural markers» (p. 42, emphases original). In equitable education environments, observable grouping patterns are nonexistent between students' background (e.g., LM status), participation measures of access, (e.g., classroom instructional activities), and outcome measures of achievement (e.g., standardized test scores). Results from the National Assessment of Education Progress (NAEP) and other standardized tests of mathematics performance, such as SAT Math, show that the sub-population of LMs score considerably lower compared to the scores of their native English-speaking peers (Abedi, 2004; Gándara & Contreras, 2009). Monitoring the between-group variation of LMs and non-LMs on standardized tests comes with warnings, such as perpetuating myths of ability, normalizing monolingualism, and expanding racial-ethnic discrimination (Gutiérrez, 2008).

A sociopolitical stance toward research is required to avoid monitoring LMs' achievement as a maintentance mechanism for reproducing deficit paradigms of teaching and learning (Gutiérrez, 2010; Valencia, 2002). Distinctive from socioculturalism, our study frames access, inclusion, fairness, and achievement as inextricable dimensions of mathematics education. Whereas sociocultural concepts include communities of practice, learning as belonging, and out-of-school mathematics, such concepts do not explicitly communicate researchers' ideological stance of social justice and transformation (Alfaro & Hernandez, 2016).

2. Methodology

With national and state high-stakes accountability requirements elevating the standardized test scores of LMs to a governance issue of critical importance, the state of California, where LMs comprise 20 percent of all public school students, cannot ignore how the language load of formative assessment processes mediate the summative standardized test scores of LMs. By conducting a content analysis (Stemler, 2001) of two textual data sources: (a) an end-of-course Algebra I formative assessment designed by a publisher and implemented by a school district in Northern California serving 20,000 students (80 percent participating in the free and reduced lunch program and over 40 percent designated as English learners), and (b) a sample CST Algebra I summative assessment available to the general public by the California Department of Education, this study specifically focused on the relationship between language load of mathematics materials on formative assessment processes and summative assessment expectations. This inquiry is guided by the following research question: To what extent is the linguistic load aligned between a formative assessment and a summative standardized test in Algebra I?

Analytic Strategy. The Algebra I California Standardized Test (CST) is a sample summative exam comprised of ninety-six questions by the California Department of Education[3] and is accessible to the general public. The Algebra I end-of-course formative assessment is comprised of thirty questions and recommended by CGP Education, a mathematics curriculum company, to determine students' readiness for subsequent coursework. Algebra I curricular materials were chosen for this study in an effort to gain a better understanding of the assessments' language load alignment using words from the Academic Word List (Coxhead, 2000). Each data source was recoded into separate word-document files without any symbolic numerical or graphical expressions. The recoded word-document files were analyzed using the same qualitative software program[4] in three stages. During stage one, the Academic Word List (AWL) content of each data source was classified into a spreadsheet with the following categories: (a) academic word and (b) context of usage. During stage two, spreadsheets were assigned preliminary descriptive codes. Guided by shared descriptive codes, data-reduction practices were used to craft intermediary analytic codes for further analysis. During the third, and final, stage of analysis, patterns of data were analyzed to determine the data sources' linguistic alignment as well as the influence of AWL terms in a summative algebraic assessment.

3 http://www.cde.ca.gov/ta/tg/sr/css05rtq.asp

4 http://www.nottingham.ac.uk/~alzsh3/acvocab/AVWL.pl

3. Results

The linguistic content of the district-adopted formative exam in Algebra I is not aligned to the summative California Standardized Test in Algebra I. For instance, the word *factor* appears once in the district-adopted test, a data source consisting of 1,030 words whereas in the CST exam, the word *factor* appears six times in a word sample of 1,722. Proportionally, the word factor comprises 0.00097 percent of the linguistic content of the district-adopted test and 0.0034 percent of the linguistic content of the CST exam. Further, the word factor is presented as an isolated and solitary directive in the district-adopted test whereas in the CST exam the word is presented across lexical categories, e.g., as a noun, «Which expression could be one of his correct factors,» as a verb, «Carter is solving this equation by factoring,» and as an adjective, «Which is the factored form of.»

Another example of linguistic misalignment in these two assessments is how the language of mathematics is employed. For instance, in the district-adopted formative assessment, the word *conclusion* appears in isolation as a multiple-choice answer option in juxtaposition with the words deduction and induction whereas in the CST exam the word conclusion appears in the following phrase, «What is the conclusion of the statement in the box below,» along with operational markers, such as «If,» «then,» and «or.» In another problem, the CST exam uses the word *conclusion* in the following context: «Which value serves as a counterexample to prove Josiah's conclusion false?» In other words, students unfamiliar with the word conclusion are offered multiple meaning-making contextual clues in the CST exam but not in the district-adopted formative assessment.

In addition to frequency and variability of word usage in assessment materials across district-adopted and standardized state exams, alignment in the academic quality of words is also paramount. Using the Academic Word List inventory developed by Coxhead (2000) as a measurement of word quality, results from this study show a disproportionate and unbalanced alignment of linguistic and content assessment across the two data sources. For instance, the district-adopted test contains 28 AWL words in a sample size of 1,030 words whereas the state exam contains 125 AWL words in a sample size of 1,722 words. Therefore, AWL words comprise 2.8 percent of the words in the district-adopted test and 7.2 percent of the CST exam. The absolute difference across the two samples is 4.4 percentage points. Such relative differences can be considered in several ways: either the district-adopted test employs AWL words 38 percent less frequently than the state assessment or the state assessment employs AWL words 257 percent more frequently than the district-adopted test.

The AWL inventory serves as a measure of linguistic alignment in another important manner: iterative frequency. Tables 1 and 2 categorize all AWL words in the two samples into three categories: (a) used once, (b) used at least twice but less than five times, and (c) used five times or more. In the state assessment AWL words only used once comprised 14 percent of all AWL words in the sample whereas in the district test, AWL words only used once comprised 64 percent of all AWL words in the sample. Conversely, while 0 percent of the AWL words were used five times or more on the district test, AWL words iterated five times or more on the state assessment comprised 65 percent of all AWL words in the sample.

Table I. Iterative Frequency of AWL vocabulary for State Algebra I Summative Assessment

AWL Frequency = 5,5 < N = 81	AWL Frequency = 2 < 5 N = 38	AWL Frequency = 1 N = 18
Equation (n = 38)	Affect, Positive, Input, & Output (n = 4)	
Equivalent (n = 10)	Conclusion, Negative (n = 3)	
Equations (n = 9)	Defined, Data, Identical, Vehicle, Area, Initial, Distributive, & Rational (n = 2)	Construction, Plus, Maximum, Exceed, Valid, Chart, Evaluated, Concluded, Prime, Region, Generate, Infinitely, Funds, Derive, Formula, Projected, Obtain, Domain
Function (n = 7)		
Parallel (n = 6)		
Factor (n = 6)		
Sum (n = 5)		

Table II. Iterative Frequency of AWL vocabulary for District Algebra I Formative Assessment

AWL Frequency = 5,5 < N = 0	AWL Frequency = 2 < 5 N = 10	AWL Frequency = 1 N = 18
		Conclusions, Deduction, Induction, Method, Infinite, Compound, Options, Undefined, Method, Function, Elements, Evaluate, Region, Coordinates, Formula, Factor, Range, Domain (n = 1)
	Equations (n = 4)	
	Volume, Dimensions, Constant (n = 2)	

4. Discussion

Mathematics educators committed to the social responsibility of enacting equitable educational policies and practices in mathematics learning environments: (a) protect students from inequitable individual, interpersonal, and institutional forms of social processes, such as discrimination; (b) promote practices that frame the participation of plural perspectives as resources for redefining constructs, and (c) enact comprehensive policies of access, inclusion, fairness, and achievement as inextricable dimensions for sociopolitical purposes. Social justice in mathematics education is evident when students make sense of data in ways that help them see the humynity behind the numbers and use mathematics as a tool for exposing and analyzing injustices in society as a means for convincing others of a particular, and often nondominant, point of view (Gutiérrez, 2010).

Attention to the language load in mathematics materials offers educators meaningful opportunities to enact equitable education policies and practices that promote inclusion and fairness for all students, and LMs in particular. Using OTL as a conceptual framework to promote inclusion and fairness for LMs., educators may consider analyzing if formative assessment processes, including district benchmarks, appropriately expose students to the content and language expectations of summative standardized tests. Moreover, educators may determine that improving students' performance on summative exams requires district investment in supportive instructional services, such as access to particular forms of technology. Finally, educators may identify a reflective process to determine to what extent their values regarding specific instructional practices are aligned with the goals of improving students' performance on summative exams.

Results from this study show, on multiple measures, that the language load of the district-adopted Algebra I formative assessment is not aligned to the summative California Standardized Test. Algebra I curricular materials selected for the analysis of lexical content offers an introduction to understanding data patterns of linguistic alignment between formative and summative algebraic assessments. Combined, data patterns between formative processes and summative standardized test scores provide a comprehensive viewpoint of what students know and are able to do as well as to what extent educators and school systems contributed to students' preparation. Considering the interdependence of language load and mathematical content knowledge and skills, school district leaders and classroom educators implementing assessment policies and practices must consider fairness and inclusion as guiding principles (Garrison & Mora, 1999). An equitable assessment program requires access to appropriate pedagogical scaffolds, including formative indicators *for* students' learning and districts will need to advocate for expert reviewers of formative assessment content (Gardner, 2006).

The linguistic alignment of districts' formative assessment materials and summative state assessments, particularly in the era of high-stakes testing, moves beyond reforming, it improves educational practices. If fostering the achievement of California students on standardized tests is important, alignment and consistency of the language load of mathematics materials in both formative assessment processes and summative standardized tests is critical—particularly for LMs. We are hopeful that future media coverage of students' achievement trends on standardized tests of mathematics will do more than communicate superficial results that perpetuate stereotypes of failure and ask policy-driven questions that set into motion an urgent call for a national formative assessment program. Specifically, we suggest journalists reporting educational assessment results explicitly communicate their sociopolitical stance and comprehensively communicate how students from minoritized backgrounds experience curricular and instructional opportunity-to-learn gaps that mediate their standardized test scores.

Bibliographic references

Abedi, Jamal (2004): "The No Child Left Behind Act and English Language Learners: Assessment and Accountability Issues". *American Educational Research Association*, 33, 4-14.

Alfaro, Critina & Hernandez, Ana M. (2016): "Ideology, Pedagogy, Access, and equity [IPAE]: A Critical Examination for Dual Language Educators". *Multilingual Educator*, March, 8-11.

Allen, Janet (2003). *Yellow brick roads: Shared and guided paths to independent reading*. Portland: Steinhouse Publishers.

Baldwin, R. Scott; Ford, Jeff C. & John E. Readence (1981): "Teaching Word Connotations: An Alternative Strategy". *Reading World*, 21, 103-108. DOI: https://doi.org/10.1080/19388078109557630.

Bullock, James O. (1994): "Literacy in the Language of Mathematics". *American Mathematical Monthly*, 10.

Carson, C. C.; Huelskamp, R. M.; & Woodall, R. D. (1993): "Perspectives on Education in America". *The Journal of Educational Research*, 86, 259-310.

Cizek, Gregory J. (2010): "An Introduction to Formative Assessment" in Heidi Andrade & Gregory Cizek (eds.): *Handbook of formative assessment*. New York: Routledge.

Coxhead, Averil (2000): "A New Academic Word List". *TESOL Quarterly*, 34 (2), 213-238.

Darder, Antonia & Mirón, Luis F. (2007): "Critical Pedagogy in a Time of Uncertainty" in N. K. Denzin & M. D. Giardina (eds.): *Contesting Empire, Global Dissent: Cultural Studies after 9/11*. Boulder, CO: Paradigm Publishers.

Darling-Hammond, Linda (2010). *The Flat World and Education: How America's Commitment to Equity will Determine our Future*. New York: Teachers College Press.

Draper, R. (2002): "School Mathematics Reform, Constructivism, and Literacy: A case for Literacy Instruction in the Reform-oriented Mathematics Classroom". *Journal of Adolescent and Adult Literacy*, 44, 520-529.

Field, Simon; Kuczera, Malgorzata; & Pont, Beatriz (2007). *No More Failures: Ten Steps to Equity in Education*. Paris: OECD Publishing.

Fisher, Douglas; Frey, Nancy, & Williams, Douglas (2002): "Seven Literacy Strategies that Work". *Educational Leadership*, 60 (3), 70-73.

Frayer, D.; Frederick, W. C.; & Klausmeier, H. J. (1969). *A Schema for Testing the Level of Cognitive Mastery*. Madison: Wisconsin Center for Educational Research.

Gardner, Howard & Hatch, Thomas (1989): "Educational Implications of the Theory of Multiple Intelligences". *Educational Researcher*, 18 (8), 4-10.

Gardner, John (2006). "Assessment for Learning: A Compelling Conceptualization" in John Gardner (ed.): *Assessment and Learning*. Thousand Oaks, CA: Sage.

Garrison, L. & Mora, J. K. (1999): "Adapting Mathematics Instruction for English Language Learners: The Language-Concept Connection" in L. Ortiz-Franco, N. G. Hernandez, & Y. De La Cruz (eds.), *Changing the Faces of Mathematics: Perspectives on Latinos*. Reston, VA: National Council of Teachers of Mathematics.

Gándara, Patricia & Contreras, Frances (2009). *The Latino Education Crisis: The Consequences of Failed Social Policies*. Cambridge, MA: Harvard University Press.

Gutiérrez, Rochelle (2010): "The Sociopolitical Turn in Mathematics Education Research". *Journal for Research in Mathematics Education*, 41.

Gutiérrez, Rochelle (2008): "A 'Gap-Gazing' Fetish in Mathematics Education? Problematizing Research on the Achievement Gap". *Journal for Research in Mathematics Education*, 39 (4), 357-364.

Gutiérrez, Rochelle (2007): "(Re)defining Equity: The Importance of a Critical Perspective" in Na'ilah Suad Nasir & Paul Cobb (eds.): *Improving Access to Mathematics: Diversity and Equity in the Classroom*. New York: Teachers College Press.

Heubert, Jay P. & Hauser, Robert M. (1999). *High Stakes: Testing for Tracking, Promotion and Graduation*. Washington, DC: The National Academies Press.

Kohn, Alfie (2000). *The Case Against Standardized Testing: Raising the Scores, Ruining the Schools*. Portsmouth, NH: Heinemann.

Lappan, Glenda; Fey, James T.; Fitzgerald, William M.; Friel, Susan N.; & Phillips, Elizabeth Difanis (2002). *Lesson Planner for Grades 6, 7, and 8*. Upper Saddle River, NJ: Prentice Hall.

Maloch, Beth (2002): "Scaffolding Student Talk: One Teacher's Role in Literature Discussion Groups". *Reading Research Quarterly*, 37 (1), 94-112.

Mehta, J. (2015): "Escaping the Shadow: A Nation at Risk and its Far-Reaching Influence". *American Educator*, 39 (2), 22-26, 43-44.

Metsisto, Diana (2005): "Reading in the Mathematics Classroom" in J. M. Kenney; E. Hancewicz; L. Heuer, Diana Metsisto & C. L. Tuttle (eds.): *Literacy Strategies for Improving Mathematics Instruction*. Alexandria, VA: ASCD.

Miller, M. David; Linn, Robert L.; & Gronlund, Norman Edward (2009): "Educational Testing and Assessment: Context, Issues and Trends" in M. David Miller Robert L., Linn, Norman Edward Gronlund (eds.): *Measurement and Assessment in Teaching* (10th edition). Columbus, OH: Merrill.

National Center for Education Statistics (2018). National Assessment of Educational Progress: An overview of NAEP. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education.

National Clearinghouse of English Language Acquisition (2011). The Growing Number of English Learner Students 09/10. Retrieved from: http://www.ncela.us/files/uploads/9/growing_EL_0910.pdf.

National Commission on Excellence in Education. (1983). A Nation at Risk: the Imperative for Educational Reform, A report to the Nation and the Secretary of Education, United States Department of Education. Washington, D.C.

National Council of Teachers of Mathematics. (2000). Principles and Standards for School Mathematics. Reston: VA: National Council of Teachers of Mathematics.

National Governors Association Center for Best Practices, & Council of Chief State School Officers (2010). Common Core State Standards (Mathematics). National Governors Association Center for Best Practices, & Council of Chief State School Officers, Washington, D. C.

OECD [Organisation for Economic Co-operation and Development], 2016). PISA 2015 Results (Volume I): Excellence and equity in education, PISA. Paris: OECD Publishing.

OECD [Organisation for Economic Co-operation and Development], (2012). Equity and quality in education: Supporting disadvantaged students and schools. OECD Publishing. Retrieved from: http://dx.doi.org/10.1787/9789264130852-en.

Provasnik, Stephen; Malley, Lydia; Stephens, Maria, Landeros, Katherine, Perkins, Robert & Tang, Judy H. (2016). *Highlights From TIMSS and TIMSS Advanced 2015: Mathematics and Science Achievement of U.S. Students in Grades 4 and 8 and in Advanced Courses at the End of High School in an International Context (NCES 2017-002)*. Washington DC.: U.S. Department of Education, National Center for Education Statistics.

Pullin, Diana C. (2008): "Assessment, Equity, and Opportunity to Learn". in Pamela A. Moss, Pullin, Diana C Pullin, James Paul. Gee, Edward H. Haertel, Lauren Jones Young (eds.): *Assessment, Equity, and Opportunity to Learn*. New York: Cambridge University Press.

Pullin, Diana C, Haertel, Edward H. (2008): "Assessment Through the Lens of 'Opportunity to Learn'" in Pamela A. Moss, Pullin, Diana C Pullin, James Paul. Gee, Edward H. Haertel, Lauren Jones Young (eds.): *Assessment, Equity, and Opportunity to Learn*. New York: Cambridge University Press.

Rodríguez, L. F., & Brown, T. (2009): "From Voice to Agency: Guiding Principles for Participatory Action Research with Youth". *New Directions in Youth Development*, 2009, 19-34.

Schwartz, Judah L. & Kenney, J. M. (1995). *Assessing Mathematical Understanding and Skills Effectively*. Cambridge, MA: Harvard Graduate School of Education.

Stack, M. (2007): "Representing School Success and Failure: Media Coverage of International Tests". *Policy Futures in Education*, 5 (1), 100-10.

Stemler, S. (2001): "An overview of content analysis". *Practical Assessment, Research & Evaluation*, 7 (1).

Tuttle, Cynthia L. (2005). "Writing in the Mathematics Classroom" in Joan M. Kenney, Euthecia Hancewicz, Loretta Heuer, Diana Metsisto & Cynthia L. Tuttle (eds.): *Literacy Strategies for Improving Mathematics Instruction*. Alexandria, VA: ASCD.

UN General Assembly (1989). Convention on the Rights of the Child. United Nations, Treaty Series, vol. 1577, p.3. Retrieved from: http://refworld.org/docid3ae6b38f0.html.

U.S. Census Bureau (2016). Language Spoken at Home by Ability to Speak English for the Population 5 Years and Over. American Community Survey 5-year Estimates. Retrieved from: datacenter.kidscount.org/data/tables/81

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (2017). Mathematics Assessments.

Valencia, Richard R. (2002): "'Mexican Americans don't value education!' On the basis of the myth, mythmaking, and debunking". *Journal of Latinos and Education*, 1 (2), 81-103.

Biogrphical notes

Saúl Isaac Maldonado is an Assistant Professor in the Department of Dual Language and English Learner Education at San Diego State University. Dr. Maldonado investigates the intersection of language, literacy and K-12 achievement using techniques such as propensity score matching and multi-level modeling on surveys, assessments, and observation protocols. Dr. Maldonado's evaluation experience includes contributions to programs and initiatives for school districts, county offices of education and federal agencies.

Rebeca Judith Andrade is Director of Early Education and Extended Learning Programs at Glendale Unified School District. Dr. Andrade researches STEM policy, practices and implementation in California public schools and project-based and social-emotional learning applications in afterschool programs. Her leadership experience includes being principal of Mountain Avenue Elementary, where she worked with staff and community to create a STEAM learning initiative, and being principal and teacher at Pasadena Unified.