

The Politics of International Large-Scale Assessment:
The Programme for International Student Assessment (PISA) and
American Education Discourse, 2000-2012

Nancy Green Sarisky

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2015

© 2015
Nancy Green Saraisky
All rights reserved.

ABSTRACT

The Politics of International Large-Scale Assessment: The Programme for International Student Assessment (PISA) and American Education Discourse, 2000-2012

Nancy Green Saraisky

The number of countries participating in large-scale international assessments has grown dramatically during the past two decades and the use of assessment results in national-level education policy debate has increased commensurately. Recent literature on the role of international assessments in education politics suggests that rankings and performance indicators can shape national educational discourse in important ways. This dissertation examines the use of one such assessment, the Programme for International Student Assessment (PISA), in education discourse in the United States from 2000 to 2012. The United States played a key role in the development of PISA and has participated in almost every international assessment of the past fifty years. Yet scholars have mostly overlooked the reception of international assessment in the United States. This dissertation seeks to address this gap.

Using an original dataset of one hundred and thirty texts from American academic literature, think tanks and the media, I examine the use of references to PISA and to top scoring countries on PISA, e.g., Finland and China (Shanghai), during the first decade of PISA testing. I find that PISA has rapidly become an accepted comparative measure of educational excellence throughout US discourse. However, despite consistently middling American scores, attempts to turn America's PISA performance into a crisis of the US education system have not stuck. Instead, I suggest that both global and domestic politics play a stronger role in shaping the interpretations of student achievement on PISA than

does student performance. I show how the American PISA discourse: (1) is driven by political, not empirical, realities; (2) contains few calls for policy borrowing from top-scoring countries and has not engendered any direct efforts at policy reform; (3) is framed with remarkable consistency across the political spectrum; and (4) is a profoundly elite enterprise, privileging the voices of international organizations and policy makers over those of parents, teachers and students.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES.....	iv
ACRONYMS.....	vi
ACKNOWLEDGMENTS.....	vii
CHAPTER ONE: Introduction	1
The Programme for International Student Assessment (PISA)	3
The United States and the Development of PISA	8
US Performance on PISA	9
Overview of the Project	10
Conclusion	15
CHAPTER TWO: Theorizing PISA Reception	17
PISA as Governance	17
Limitations to PISA as Governance	21
PISA in the Media.....	23
Limitations to PISA in the Media	24
PISA Reception as Evidence Usage	25
Performance Data as Evidence	
Indicators and the Policy Process	
PISA and the Policy Process	
Political Discourse and Agenda Setting	
Externalization and referencing	
Conceptualizing PISA Reception in the United States.....	35
The Education Policy Context in the US, 1983-2012.....	36
Comparison as Problematization	
Accountability and Assessment	
The Tempering of Partisanship	
Shifts in the Locus of Policy Activity	
The Development of the Common Core State Standards (CCSS) and the Use of PISA	
Conclusion	47
CHAPTER THREE: Using Content Analysis of Educational Discourse to Understand PISA Reception in the United States.....	48
Research Design	48
American Policy Knowledge	50
Academics and peer-reviewed literature	
Policy analysts and think tank literature	
Journalists and the media	
Sampling	55
American academic literature	
Think tank literature	

American media	
Exclusion rules	
Overall sample	
Data Analysis	66
Academic and think tank literature analysis	
Media analysis	
Limitations	73
 CHAPTER FOUR: What we talk about when we talk about PISA:	
Findings from American educational discourse	76
Overview: PISA across literatures, 2000-2012.....	79
Academic literature	82
PISA as a Data Source	
Limits to PISA data	
PISA as a Model of Assessment	
PISA as a Measure of Educational Excellence	
Academic Use of Reference Societies	
No Lesson Drawing or Calls for Emulation	
Conclusion	
Think tank literature	105
Think Tank References to PISA Over Time	
Traction through “Crisis Talk”	
The Rhetoric of Sputnik and China	
PISA as a Measure of American Educational Excellence	
PISA as a Measure of Foreign Educational Excellence	
Think Tank Critique of PISA	
Think Tank Rhetoric	
Think Tank Politics: Linking PISA to the Common Core State Standards	
Different Ideologies, Similar Claims	
Think Tank Use of Reference Societies	
Minimal Calls for Lesson Drawing	
Conclusion	
Media	131
From “Good News, Bad News” to “Stunned Educators”: Two Phases of	
Discourse about PISA	
Media Use of Reference Societies and Crisis Talk	
Explanations of Finnish success	
Explanations of Chinese success	
Framing of PISA in US Discourse	
Critique of Using PISA as Evidence of a Need for Improvement	
Voice in PISA Discourse	
Elites as Experts	
The OECD and the media	
Role of opinion writing	
PISA as a canvas for education elites	
Conclusion	

CHAPTER FIVE: PISA in comparative perspective: Reception in Germany, Japan and the United States	
Comparative Case Study Methodology	184
Limitations	186
Capturing the Discursive Effects of PISA in Three Countries	187
The quantity of media coverage	
The tone of media coverage	
The effects of PISA on policy	
Reception in Germany, Japan and the United States, 2000-2012	189
Germany	
Japan	
The United States	
What Explains Variation in PISA Reception?	200
Scores	
Timing	
Actors	
Federalism	
Problem Definition	
Conclusion	210
CHAPTER SIX: The Politics of PISA Discourse in the United States, 2000-2012	214
Discussion of Findings.....	214
Importance of Timing	
Cleavages between Literatures	
Discourse is Driven by Politics, not Evidence	
Use of References to Top Scorers	
Additional Puzzles	226
PISA as Ranked Data	
PISA and Politics	
PISA in Discourse	
Areas for Future Work	234
Indifference as a Response.....	235
Conclusion	236
REFERENCES	239
APPENDIX A: SAMPLED TEXTS	258

LIST OF TABLES AND FIGURES

- Figure 1.1: Country/Economy Participation in PISA, 2000-2009
- Figure 1.2: Example of a PISA league table, 2009 Results
- Table 1.3: US Performance on PISA, 2000-2009
- Figure 3.1: Model of Policy Knowledge Production
- Table 3.2: American Education Policy Knowledge Production
- Table 3.3: Overview of AERA Journal Sample
- Table 3.4: Overview of Think Tank Sample
- Table 3.5: Overview of Media Sample
- Figure 3.6: Sampled Proportion of Articles/Reports to Blog Posts in Think Tanks and Media
- Figure 4.1: References to PISA in academic, think tank and media literatures, 2000-2012
- Figure 4.2: Cumulative references to PISA, 2000-2012
- Figure 4.3: Overview of academic sample, by journal
- Figure 4.4: Academic journal articles, by PISA frame
- Figure 4.5: Use of reference societies in academic literature
- Figure 4.6: Think tank texts by frame
- Figure 4.7: Think tank references to Finland and China
- Figure 4.8: Think tank explanations of Finnish and Chinese success
- Figure 4.9: Overview of media sample, 2000-2012
- Figure 4.10: Placement of PISA coverage, 2000-2012
- Figure 4.11: Coverage of PISA by discourse moment and by source
- Figure 4.12: Media references to Finland and China, 2000-2012

- Figure 4.13: References to Finland and China by media source
- Figure 4.14: Media explanations of Finnish and Chinese success
- Figure 4.15: Media frames by use (percentage)
- Figure 4.16: Overview of speakers, by affiliation
- Figure 4.17: Academic speakers and university affiliations
- Figure 4.18: Speakers by number of speech acts
- Figure 4.19: Speech acts by speaker affiliation and source
- Figure 4.20: Examples of PISA projections
- Table 5.1: PISA reception in Germany, Japan and the United States, 2000-2009
- Table 5.2: PISA performance and rank (in parentheses): Germany, Japan and the United States, 2000-2009

ACRONYMS

CCSS	Common Core State Standards
CCSSO	The Council of Chief State School Officers
CIVED	Civic Education Study
EC	European Commission
IEA	International Association for the Evaluation of Educational Achievement
ILSA	International large-scale assessment
NAEP	National Assessment of Educational Progress
NAR	A Nation at Risk
NCLB	No Child Left Behind
OECD	Organisation for Economic Development and Cooperation
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
RttT	Race to the Top
TIMSS	Trends in International Mathematics and Science Study

ACKNOWLEDGMENTS

Many people have guided and supported me in the completion of this dissertation. My first thanks go to my stellar dissertation committee. My committee chair, Gita Steiner-Khamisi, has been an adviser, friend and supporter since my very first days at Teachers College, when we bonded while discussing our experiences in Mongolia. Gita has been endlessly supportive and patient throughout my doctoral journey and has offered intellectual and conceptual guidance that was critical to the project. Jeff Henig has been instrumental in my thinking. In his doctoral seminar and beyond, his comments forced me to consider whether I was asking the right questions, and then how to best discover the right answers. Ken Prewitt's work on data use helped me to consider the role of international assessment as "evidence-influenced politics," and his feedback has been insightful and helpful. I am grateful for his participation, especially making time to participate in the defense from a distance. Tom Hatch also provided useful comments on the research design and other aspects of the dissertation, for which I am most appreciative.

I owe a special debt of gratitude to my sponsor, Oren Pizmony-Levy. From the moment I first spoke with Oren (before he had even officially arrived at TC), he has been unflinchingly supportive of my work. Our dozens of hours of discussion about international assessment have been intellectually stimulating, enormously influential and a lot of fun. Oren immeasurably improved my work theoretically, empirically, and interpretively. He is also a model of what a mentor and colleague can be. He deserves full credit for making the writing of this dissertation an outstanding experience.

For financial support, I owe thanks to the National Science Foundation's Graduate Research Fellowship and to Teachers College. Also from Teachers College, I wish to thank Gary Ardan, who has been so helpful in solving problems and making all the administrative requirements of being a doctoral student bearable. The department of International and Transcultural Studies at TC has been a good home; in particular, Fran Vavrus, Lesley Bartlett, Dana Burde, Garnett Russell and Carol Benson all offered helpful advice and support along the way.

Many friends helped to make my doctoral studies exciting and stimulating, but Julieta Garcia-Hamilton, Constanza LaFuente, Eleni Natsiopoulou, and Kyle Long deserve special mention for keeping me laughing and well-cafeinated, and providing feedback whenever asked.

Outside of TC, I am blessed with the boundless support of a wonderful family. Susan Green was, and is, always there, no matter what, with a smile and usually with something good to eat, too. The late Edward Green taught me to be curious, and I know he is proud that he finally has a doctor in the family. Ellen and Al Saraisky never stopped asking how it was going and were an endless source of encouragement. Sophie, Madeleine, Hannah, Lucy, Lauren, Donna, Cindi, Richard and many other relatives were perpetually caring and encouraging. And finally, my love and gratitude go to Steve, who supported me with patience, humor and love, as he always does.

Chapter One: Introduction

During the past three decades, international large-scale assessments (ILSAs) of student achievement have increasingly become litmus tests by which national educational systems are judged. As the number of countries participating in ILSAs has grown (Kamens and McNeely, 2010; Kamens and Benavot, 2011; Pizmony-Levy, 2013), the use of assessment results in national-level education policy debate has increased commensurately. Often citing the demands of a “knowledge-based economy” and the necessity of national economic competitiveness in a globalized economic world, governments turn to the results of international assessments with such frequency that it has become almost standard practice (Steiner-Khamsi, 2010; Pizmony-Levy, 2014; Howie and Plomp, 2005).

The use of ILSA performance data in policy debate most commonly involves referring to country performance from a ranked list, where countries are arranged in a list according to achievement scores. These rankings are powerful ways of disseminating information as they simplify reality by making all countries and education systems comparable, regardless of underlying economic, political or systemic characteristics. If education policy makers know they are being monitored and ranked, they may be more apt to promote policy reform to meet expectations set by the rankings (Espeland and Sauder, 2007). In fact, recent literature has argued that rank and performance indicators are

impacting national education policy (Breakspear, 2012; Wiseman, 2013; Martens and Niemann, 2010) and discourse (e.g., Takayama, 2008; 2010) in important ways.

Thus ILSAs are a potentially important source of pressure and information in national education politics and policy. National policy makers, ever concerned about how their education systems measure up, can easily use ILSA rankings to benchmark their country's performance against the competition. Local media can highlight the drama of storylines about which countries are top scorers. And the public is offered a simple numerical snapshot of its domestic education system. Understandably, these assessments are seen to have an important and growing impact on education policy worldwide.

Steiner Khamsi (2003) categorized the impact of ILSA scores in a three-part typology: scandalization (highlighting the weaknesses of the home country as a result of comparison); glorification (highlighting the strengths of the home country as a result of comparison); and indifference, which Steiner-Khamsi leaves undefined. A great amount of attention has been paid to narratives associated with scandalization and "PISA shock," where scores appear to have thrown the national education environment into turmoil, as in Germany in the early 2000s. But a closer reading of the extant literature shows that the impact of ILSA rankings may be more complex than these crisis narratives suggest (e.g., Grek, 2009; Pons, 2012). While crisis and scandalization are dramatic, reactions to education performance rankings are far from uniform. Little empirical work has been done to account for variation in ILSA reception, or to look systematically at the

mechanisms that link international assessment to domestic education policy and politics.

This dissertation attempts to begin to fill this gap.

To better understand the interaction between ILSAs and national education policy, I explore the reception and use of one such assessment in one country context.

Specifically, I examine the use of the Programme for International Student Assessment (PISA) in American educational discourse from 2000-2012. The picture that emerges offers both a more systematic and more nuanced account of PISA's effect on domestic politics, and the role domestic politics play in PISA reception.

The Programme for International Student Assessment (PISA)

PISA is an international assessment coordinated by the Organisation for Economic Cooperation and Development (OECD) and administered every three years to a representative sample of 15-year old students from participating countries.¹ Participation is open to OECD member states as well as non-member states and certain sub-national units.² PISA measures student performance in three subject areas: mathematics, reading and science literacy. Each testing cycle covers material from each of these subject areas but focuses in-depth on one subject. The subject of focus rotates with each cycle (i.e., PISA 2000 focused on reading literacy, 2003 on mathematics, 2006 on science, and so

¹ PISA employs probability proportional to size (PPS) sampling, a multistep process that involves sampling schools and then sampling eligible 15-year olds from within the selected schools. For a description of the sampling methodology, see <https://nces.ed.gov/surveys/pisa/faq.asp>.

² For example, two states in India (Tamil Nadu and Himachal Pradesh) participated in PISA 2009, and Hong Kong, Macau and Shanghai each participate in PISA as separate units rather than as "China."

on). Students who sit for PISA also complete a lengthy background questionnaire, which is administered to school principals as well.³ Results are officially released in early December the year after the assessment is given; that is, PISA 2000 results were released in December 2001, PISA 2003 results were released in December 2004, and so forth.

Since it was first administered in 2000, PISA has arguably become the most prominent international assessment. Forty-three countries took part in the first round of PISA testing in 2000. By 2009, the test had 75 countries/economies participating (Figure 1.1). Initial participants were primarily OECD member countries, but participation has grown significantly and in 2009 more than half of the participants were non-OECD members.⁴ PISA is now the flagship activity of the OECD's Education Directorate, and it accounts for thirty per cent of unit's budget (Grek, 2009, 25).

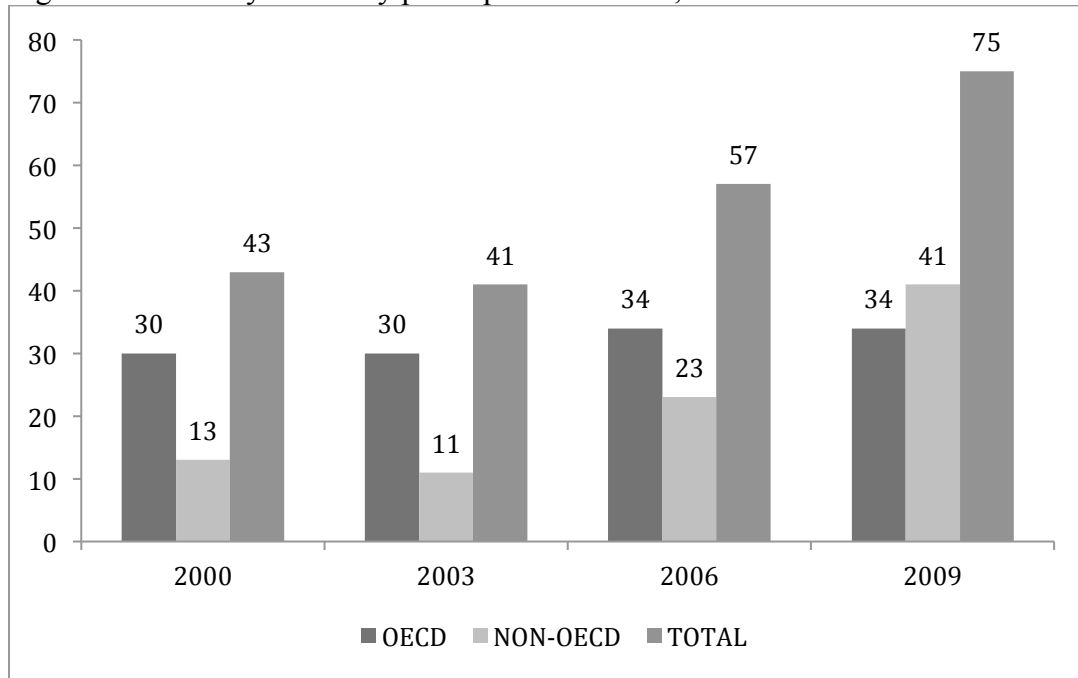
PISA differs from other international assessments in both content and purpose. Most notably, PISA does not test material that is directly tied to any school curricula. Instead, PISA is designed to “assess to what extent students at the end of compulsory education can apply their knowledge to real-life situations and be equipped for full participation in society” (OECD, 2014a). This focus on problem solving, knowledge application and functional skill differs from other well-known international assessments, such as Trends in International Mathematics and Science Study (TIMSS), administered by the

³ Beginning in the 2015 cycle, a sample of teachers will also complete the background questionnaire.

⁴ See Meyer and Benavot (2013); Kamens and Benavot (2010); Kamens and McNeely (2010); Kamens (2013); and Lockheed (2013) for empirical discussions of this phenomenon. Countries must register to participate in PISA well in advance (registration for the 2018 cycle was closed at the end of 2014) and the OECD anticipates 75 participants in PISA 2015.

International Association for the Evaluation of Educational Achievement (IEA), or national assessments such as the National Assessment of Educational Progress (NAEP) in the United States. TIMSS and NAEP are designed to align with and examine mastery of curricular knowledge.⁵ PISA instead evaluates students on their application of knowledge to ‘real world’ problems and situations.

Figure 1.1: Country/Economy participation in PISA, 2000-2009



Additionally, PISA is marketed specifically as a policy tool. The OECD promotes PISA as an instrument that can “reveal what is possible in education” (OECD, 2014a) by

⁵ That being said, country results on PISA and international assessments that do align to national curricula, such as TIMSS, are generally highly correlated.

providing information about high-performing and rapidly improving systems.⁶ The OECD describes the assessment as allowing policy makers to

“use PISA results to gauge the knowledge and skills of students in their own countries in comparison with those in other countries, set policy targets against measurable goals achieved by other education systems, and learn from policies and practices applied elsewhere” (OECD, 2014a).

This is in contrast to other international assessment studies that are designed to test curricular knowledge. These assessments, such as TIMSS, PIRLS (Progress in International Reading Literacy Study) and others administered by the IEA, have their roots in an intellectual orientation to understand the similarities and differences of national education systems, rather than in a political orientation to prescribe policy.⁷ The IEA assessments were initially designed to investigate country differences not through ranking but by acknowledging each country’s unique educational and cultural histories (Kandel, 1959; Kamens and McNeely, 2010). This is markedly different from PISA,

⁶ For instance, beyond the usual technical reports and academic analyses of PISA data, the OECD website has a page for “PISA Products.” This includes a series of reports and videos produced in conjunction with the Pearson Foundation called “Strong Performers and Successful Reformers in Education” that highlights best practices of top-scoring countries. See <http://www.oecd.org/pisa/pisaproducts/>

⁷ As Pizmony-Levy (2014) recounts, the IEA was started by a group of scholars, under the auspices of UNESCO, to collaboratively investigate specific research concerns. The individuals involved were academics or researchers, and developed an agenda around specific scholarly research questions. The minutes from a meeting of the IEA describes the goals of the early assessments this way: “The aim is to develop a systematic study of educational outcomes in the school systems of the cooperating countries. The question we wish to ask is not “Are the children of country X better educated than those of country Y?” To us this seemed a false question begging all the important issues we need to study” (Minutes of the IEA Project, in Pizmony-Levy, 2014).

where contextual and cultural differences are erased so that country scores can be compared and interpreted as “apples-to-apples.”

A third defining characteristic of PISA has been the use of its scores to create ranking tables, sometimes called league tables (see Figure 1.2). In line with the modern privileging of scientific rationality in the form of numbers and indicators, league tables have become a prominent and persuasive form of evidence (Pons and van Zanten, 2007; Mehta, 2011; Davis, Kingsbury and Merry, 2012). These tables provide simplified, rank-ordered data that is designed to set a standard for evaluating performance. The tables visualize a process that sociologists call commensuration; this process collapses any differences between units and allows countries to be compared on a single, common metric (Espeland and Stevens, 1998). It is impossible to know from the looking at the education league tables whether the differences between country scores are significant or not, or whether the scores show a change from a previous round of testing, or what factors account for any changes. Instead, the ranked lists encourage a simplistic understanding of winners (those on the top of the tables) and losers (those at the bottom).⁸ The ranking tables also become the basis for benchmarking, where one education system

⁸ The decontextualized nature of the tables encourages at least three types of analyses about which there should be caution. First, the decontextualization of data provides local actors with a blank slate for interpretation; success, or failure can be attributed to multiple causes. Second, the tables invite the temptation to select on the dependent variable. Top-scoring countries – those that are ‘successful’ – are the ones to be analyzed, rather than analyzing countries at the top, middle and bottom. This dovetails with another problem that Loveless calls the “A+ country fallacy.” Analyses focus on the characteristics of top-scoring countries and conclude that those characteristics must account for success. For instance, many analyses cite the fact most top-scoring countries have a national curriculum. Alas, most low scoring countries have a national curriculum as well. See Loveless (2012).

measures its quality and performance by comparing itself to the performance and ‘best practices’ of its peers.

Figure 1.2: Example of a PISA League Table, PISA 2009 results

Mathematics			Science			Reading		
1	 Shanghai, China	600	1	 Shanghai, China	575	1	 Shanghai, China	556
2	 Singapore	562	2	 Finland	554	2	 South Korea	539
3	 Hong Kong, China	555	3	 Hong Kong, China	549	3	 Finland	536
4	 South Korea	546	4	 Singapore	542	4	 Hong Kong, China	533
5	 Taiwan	543	5	 Japan	539	5	 Singapore	526
6	 Finland	541	6	 South Korea	538	6	 Canada	524
7	 Liechtenstein	536	7	 New Zealand	532	7	 New Zealand	521
8	 Switzerland	534	8	 Canada	529	8	 Japan	520
9	 Japan	529	9	 Estonia	528	9	 Australia	515
10	 Canada	527	10	 Australia	527	10	 Netherlands	508

Source: Max Planck Institute for Human Development 2015

The United States and the Development of PISA

The United States was critical to the early development of the OECD⁹ and of PISA. The OECD, created as part of the post World War II reconstruction architecture, had no mandate for education specifically. But insofar as education was seen to be a component of economic growth and international competitiveness, education was on the OECD agenda from the beginning. However it was only in the mid-1980s, thanks to a major push from the United States, that the education work of the OECD began to focus extensively on measurement. This was in spite of extensive opposition from other member countries and OECD staff. As Stephen Heyneman describes:

“The year after the publication of *A Nation at Risk*, I visited OECD headquarters... The US delegate was said to have put a great deal of pressure, and

⁹ See Henry, Lingard, Rizvi and Taylor (2001) for a detailed treatment.

in very direct language, for the OECD to engage itself in a project collecting and analyzing statistical education “inputs and outcomes” – information on curricular standards, costs and sources of finance, learning achievements on common subject matter, employment trends and the like. The reaction among the staff at CERI (the Center for Education Research and Innovation) was one of shock, and deep suspicion. Those whom I interviewed believed it was unprofessional to try and quantify such indicators, that it would oversimplify and misrepresent OECD educational systems, and that it would be rejected by the 24 member states whose common interests they were charged to serve. More important, they believed that the demand for such information would shift as soon as a Democrat could come to the White House” (Heyneman, 1993, 375).

Heyneman’s account is a fascinating set-up for the story to come. At the time, conservative politicians in the Reagan Administration were pushing to abolish the US Department of Education completely. The OECD strongly resisted the initial development of a PISA instrument because it ran counter to the OECD’s understanding of country systems as entities that were separate and different. As with the designers of the early IEA comparative studies, the OECD viewpoint on this issue was that of a research institution. The OECD had not deeply concerned itself with the development of education policy, or as a purveyor of educational best practice. However, as I discuss in more detail at the end of this chapter, the US became extremely interested in international comparative data in the early 1980s, an interest that only increased after the publication of *A Nation at Risk (NAR)* in 1983. In the next chapter I discuss the impact of *NAR* as a turning point in American education discourse more fully. Here it is simply to note that the US was the prime mover behind the development of PISA at a time when the OECD had very little interest in developing the instrument.

US Performance on PISA

US performance on PISA has remained fairly constant – with scores at or below the OECD average – since the test began more than a decade ago (Figure 1.3). This is in line with American performance on international assessments in general since the 1960s, on which American performance has been middling at best. These results have generally held across tests and over time. America’s performance has in the past led to objections about the utility and validity of such comparisons. As Stedman (1994) points out, critics have raised doubts about sampling bias (samples were not representative), test bias (curricular or cultural), and validity of the tests again and again as the US has underperformed (Mehta, 2006). Despite its consistently mediocre performance, the US has remained a strong supporter of international assessment.

Table 1.3: US PISA performance over time

Subject	Country	2000	2003	2006	2009
Mathematics*	OECD	493	500	494	496
	United States	493	483**	474**	487**
Reading	OECD	494	494	489	493
	United States	504	495	---	500
Science	OECD	494	499	498	501
	United States	499	491**	489**	502
<p>* Significance tests were not available for 2000 mathematics scores.</p> <p>** Score is statistically significantly lower than the OECD score, which represents the international average.</p>					

Overview of the Project

Though the United States played a key role in the development of PISA and has participated in almost every international assessment of the past fifty years, scholars have mostly overlooked the reception of PISA in the US. My dissertation seeks to address this gap by conducting a systematic analysis of PISA reception in the United States throughout the first decade of PISA testing. The dissertation focuses on three main questions: how has PISA been received and understood in US education discourse between 2000 and 2012? Who are the actors contributing to PISA reception in local discourse? How does PISA reception in the US vary across actors, reflecting agency and interests?

My research shows that while PISA has become a taken-for-granted form of evidence in American educational discourse, its use varies over time. PISA's reception and use in the United States is highly contextually dependent, driven by geo/political interests exclusive to the United States. Far from being a new form of governance eclipsing the state, as some have claimed, PISA appears to have had minimal effect on US education policy.

As I demonstrate in the empirical chapters, the US is an important and paradoxical case. Although the US holds a deep belief in the importance of standardized testing, and was pivotal in the development of PISA, PISA has not been used to obvious effect as evidence for educational policy change in the US. Although US scores on PISA have remained remarkably consistent over time, the American reaction to those scores has varied wildly. And despite the fact that PISA is linked to an accountability regime that is

supposed to provide evidence about the performance of teachers and students, teachers and students have essentially no voice in the discourse about PISA.

I find that while PISA has rapidly become a commonly referenced measure of educational excellence in US discourse, attempts to turn America's PISA scores into a crisis of the education system have not stuck. Nor has the United States' middling performance on PISA resulted in calls for policy borrowing from top-scoring countries. Instead, I show how the American PISA discourse: (1) changes over time; (2) is driven by political concerns, not evidence; (3) contains few calls for lesson drawing or policy borrowing; (4) is framed with remarkable consistency across the political spectrum; and (5) is a profoundly elite enterprise.

The research extends the understanding of PISA reception theoretically. Whereas many existing studies of PISA reception focus primarily on the use of PISA scores to scandalize the home education system and compel policy change, I find that the US case is more nuanced and does not fit easily alongside existing accounts. After a contained period of crisis talk, the US proves to be rather indifferent to PISA results vis-à-vis domestic education policy. This is contrary to existing studies of international assessment that argue that references to PISA results and to top scorers are made to legitimize contentious policy in the home country. There is little evidence in the American discourse that policy actors are using PISA to justify contentious policy action,

such as the implementation of the Common Core State Standards (CCSS) that began in the late 2000s.

To explore PISA reception in the US, I draw on ideas from political science and political sociology about performance indicators and their influence on behavior; from the policy studies literature that discusses agenda setting; and from an interpretive framework in comparative education. Performance indicators are theorized to induce a change in behavior as a result of the pressure they can exert on policy makers (Espeland and Sauder, 2007; Kelley and Simmons, 2014). Vis-à-vis policy, indicators function as an agenda setting mechanism, by helping to identify and name public problems, create causal theories about responsibility and blame, and by constructing solutions to those problems (Gusfield, 1981; Stone, 1997; Mehta, 2011). If problems and solutions are not clearly defined, however, performance data may be less politically useful. In the United States, PISA discourse does not problematize American education in a way that has been useful for policy elites. Consequently, actors and interests have not come together and PISA has not had a disruptive effect on policy.

In a preliminary attempt to assess under what conditions PISA is used to affect policy, I then compare the US case to reception in Germany and Japan, where the political reaction has been full-blown “PISA shock.” Drawing on extant literature about Germany and Japan, I begin to identify some factors that might influence PISA across national

contexts. The comparison shows that problem definitions are relevant but they must be connected to pre-existing policy activity to gain the most leverage.

My approach extends recent scholarship in at least four ways. First, I add a new case to the literature. Despite the fact that the United States played a key role in the development of PISA and has participated in almost every international assessment of the past fifty years, it has been largely overlooked in scholarship. The existing research (see Bieber and Martens, 2011; Martens and Niemann, 2010) does not systematically pay attention to whether and how PISA scores have been used in American education politics. I provide a deeper analysis and more thorough discussion of the United States case than previously available.

Second, I broaden the sampling parameters for collecting data, resulting in a more robust dataset than other scholars have used previously. I go beyond analyzing references to PISA only in the media. While the literature has long demonstrated the critical role media play in political agenda setting (e.g., McCombs and Shaw, 1972; Gamson and Modigliani, 1989; Chong and Druckman, 2007; 2011), I extend the venues in which the creation of political discourse occurs. I use academic and think tank literature to broaden the scope of actors and institutions that construct the discursive space. As I discuss in Chapter Three, I also employ a more systematic method of data analysis.

Third, unlike most of the existing literature, I explore the development of discourse over more than a decade. Existing studies of PISA reception tend to focus on reception as somewhat static, occurring either in response to the release of data from a specific PISA cycle or during a limited timeframe.¹⁰ Yet a more thorough and complex picture can be drawn from data collected over a longer period of time. I sample data from the first four cycles of PISA corresponding to the first decade of testing. While this twelve-year period is arguably historically shortsighted, representing one snapshot in time, it is nonetheless a longer time horizon than in previous studies. This allows for a richer examination of the discourse around PISA, including the comparison to “top-scorers,” e.g., Finland and China, over time.

Finally, whereas existing research tends to focus on single-case studies, I incorporate a limited comparative account of PISA reception, comparing the US case to accounts of PISA reception in Germany and in Japan. This allows me to begin to assess under what conditions PISA data are used in national education discourse, and what variables might influence how PISA is received in different national contexts.

Conclusion

Though the OECD was uninterested when the US first approached it about developing PISA, US pressure eventually won out. PISA is now widely regarded as playing a key role in international education policy and discourse. To further examine the role of PISA,

¹⁰ In fairness, some of the work that looks at only one or two cycles of PISA has to do with the fact that those were the only data available at the time of writing.

I explore PISA reception in the United States from 2000-2012. The dissertation proceeds as follows. In Chapter Two I review the literature on PISA reception and provide a theoretical framework. Bringing together theory from political science, sociology and comparative education, I develop synthetic approach that uses discourse as a way to elucidate the interplay of actors, agency, self-interest and ideas that contribute to PISA reception. A central argument here is that the development of understandings and discourse about PISA cannot be divorced from, and in fact reflect, broader trends and ideas in US education politics. It is the combination of specific American contextual factors that shapes when, how and why PISA has come to be understood as it has. Thus in Chapter Two I also briefly discuss a few trends in the American education landscape that are key to understanding reception. I also introduce the implementation of the Common Core State Standards as way to apply theory to the US case. In Chapter Three, I discuss the research design, including sampling, data collection, methods and limitations. Chapters Four and Five present empirical findings. Chapter Six offers a discussion and concludes.

Chapter Two: Theorizing PISA Reception

Within the broader field of research that looks at policy borrowing and lending, a recent and growing line of work has focused on international assessments as its object of study. These international assessment studies examine the implications of the recent exponential growth of ILSAs, in both assessment type and assessment participation (Kamens and Benavot, 2011; Kamens, 2013; Kamens and McNeely, 2010; Lockheed, 2013) by looking at how ILSAs are understood in various country contexts. This scholarship aims to map specifically how international assessment results are used locally, and “how they are locally modified and implemented ... and what impact they have on existing structures, policies and practices” (Steiner-Khamisi, 2014, 162). Given its rising global profile, it is not surprising that PISA is the subject of many of these reception studies.

PISA as governance

The scholarship on PISA reception can be categorized into two broad groups. The first understands PISA as a new form of governance. These studies identify a similar ‘PISA effect’ in many different country contexts, highlighting PISA as a “new mode of global education governance in which state sovereignty over educational matters is replaced by the influence of large-scale international organizations” (Meyer and Benavot, 2013, 10).

In this work, authors argue that PISA has shifted power from the national structures of education to the international results of the test, which are then used to govern and affect

policy in national contexts. Implicitly, these scholars are underscoring the sociological process of commensuration, whereby complex phenomena are simplified into a single data point in order to construct a space for comparison. Commensuration is powerful because it quantifies, simplifies, rationalizes and standardizes (Espeland and Stevens, 1998; 2008; Davis, Kingsbury and Engle Merry, 2012). Performance metrics are particularly powerful because they are quantitative and systematic, and thus rational, scientific, and authoritative (Desrosières, 1998; Kamens and McNeely, 2010; Henig, 2008a). They are also assumed to wield a kind of soft power that, through political and social pressure, can result in a change in behavior (Espeland and Sauder, 2007).

With regard to PISA, the commensuration process allows for any state education system to be compared to another, regardless of the myriad differences that inform local constellations of power, politics and policy (Coffield, 2012; Grek, 2009; Lingard and Rawolle, 2011; Sellar and Lingard, 2013a). Comparison of PISA performance data becomes key to defining educational success. Education systems “cease to be seen as unique and everything can be compared” (Kamens, 2011, 122). These scholars emphasize the culture of comparison of which PISA is emblematic (e.g., Nóvoa and Yariv-Mashal, 2003; Ozga, 2012; Grek, 2009).

The emphasis on PISA as governance is especially vivid in literature from and about Europe. Grek (2009) argues that PISA has become “an influential tool of the new political technology of governing the European education space by numbers” (p. 23).

Grek, Lawn, Lingard, Ozga, Rinne, Segerholm, et al., (2009) suggest that data, most prominently in the form of PISA results, are “operating to shape and influence education policy in the European systems explored here” (p. 13; the countries they analyzed were England, Sweden, Finland and Scotland). Sellar and Lingard (2013b), though focusing on Australia, also stress that the development of tools that provide comparative statistical data and allow for a process of commensuration. The argument is that PISA creates a global education policy space by existing as a common instrument through which disparate countries can be measured. That space is then governed by the very same data that created it. PISA is seen as a key regulatory tool in this regard (Carvalho, 2012; Ozga, 2012).

Though these scholars conclude that PISA is governing this global education policy space, a close reading of the data seems to offer a different conclusion. The narrative is less clear when it comes to the actual effects of PISA on national education policy. For instance, as part of the European Union’s KnowandPol project, which aims to understand the “fabrication, circulation and use of knowledge regulation tools” (www.knowandpol.eu), research teams in six countries used document analysis, interviews and media reviews to look at PISA reception. Bajoni, Berényi, Neumann and Vida (2009) conduct semi-structured interviews and a media analysis on PISA reception in Hungary. They find that PISA results are used to focus existing discourses in education but not to produce systemic policy change. In Portugal, researchers find that PISA is heavily referenced in public discourse but do not detail any policy change based

on PISA data (Afonso and Costa, 2009, 88-89). In Romania, the research team finds “no detectable relationship between the PISA results and the developing reform processes” (Róstas, et al., 2009, 35).

Grek (2009) looks at PISA reception in Finland, Germany and the UK. Her data suggest that PISA scores have had little policy impact. Grek discusses the fact that regardless of Finland’s top performance on PISA, the Finnish Ministry of Education went ahead with longstanding reform plans anyway (p. 29). In Germany, there was a ‘PISA shock’ but PISA results were used to fast-track reforms already well underway. In the United Kingdom, she reports “no concrete initiatives were undertaken in the UK – in either England or Scotland – in response to PISA results” (Grek, 2009, 31). Grek also reports on interviews with two European policy-makers in Brussels at the European Commission (EC). These informants report that while PISA data are ubiquitous and of high quality, their greatest effect seems to have been in spotlighting some gentle competition between the EC and the OECD. The interviewees say that PISA has pushed the EC to develop its own capacity for data collection. This suggests a change in behavior rather than a change in policy.

In another study of the policy impact of PISA, Breakspear (2012) surveys members of the PISA Governing Board¹¹ to assess PISA’s impact on national education policy and

¹¹ Breakspear sent questionnaires to the 65 members of the PISA Governing Board, who are each appointed to serve on the Board by their national ministry of education. Board members are mostly affiliated with

practice. Though 31 out of 37 respondents said that PISA was extremely, very, or moderately influential in informing the policy making process at the national level, respondents reported few specific policy reforms that were actually linked to PISA. Instead, respondents simply agreed that PISA rankings were an important indicator of system performance.

While making the case that in global education policy and politics the authority has shifted to supranational levels, Lingard and Rawolle (2011) show that OECD data was used by the media to affect discourse in Australia, though they do not link this discourse to specific policy. Wiseman (2013) argues that PISA has led policy makers to talk about similar education issues in light of PISA scores. He notes, however, that PISA does not actually promote similar answers or solutions to these issues. Wiseman's work suggests that PISA rankings are used not for policy change but rather that PISA has become a part of the larger world culture of education.

Limitations to PISA as Governance

Despite claims that PISA now governs national policy spaces, the limited empirical data seems inconclusive at best. While the literature argues that PISA is a new form of governance, it provides scant data in terms of PISA's actual policy impact or as a policy-transfer mechanism. PISA is clearly an important new source of educational knowledge. What is less clear is if and how this knowledge is getting translated into educational

Ministries of Education, though occasionally members are listed as being affiliated with an academic or research institution. See <http://www.oecd.org/pisa/contacts/pisagoverningboard.htm>.

policy or why, for instance, similar PISA scores trigger dissimilar policy responses on the ground (Meyer and Benavot, 2013, 15).

In addition, research that focuses on PISA as a mode of governance downplays the agency of states, institutions and actors in using PISA to define or manipulate domestic politics. It also emphasizes a similarity of experience, rather than the unique local contexts that enable and constrain the interpretation of PISA on the ground. Even with the huge amount of influence the OECD wields in terms of promoting and marketing PISA and its products, national level governments still control the power to participate in PISA. To paraphrase Ernst Haas from a slightly different context, it is states, not international organizations, which are the architects of national education policy (Haas, 1990, 6).

It is clear that the power of ILSA data – even when it is consistently simplified, decontextualized, and ripe for interpretation -- is not monolithic. PISA has undoubtedly exponentially enhanced the OECD's role as a source of educational authority in Europe and beyond, despite the fact that institutionally the OECD has neither the budget nor the legal instruments to implement policy at the national level (e.g., Grek, 2009; Papadopoulos, 2011). While there is no doubt that the geography, politics and culture of Europe put it in a different relationship to PISA vis-à-vis the OECD, the pan-European processes and organizations that are engaged in educational policy discourse do not have veto power when it comes to national education policy. National education policy is still

primarily a function of the state. PISA ‘governance’ is a matter of persuasion on the strength of ideas (Béland and Orenstein, 2012). In this regard, PISA is perhaps less a new form of governance than a new form of evidence.

PISA Reception in the Media

The other set of PISA studies moves away from the idea of governance and focuses specifically on local impact. These studies take media as the site of reception and, not surprisingly, find that reception varies across countries. These studies inherently acknowledge the role that media play in framing public issues, and influencing both elite and public opinion.

For instance, Martens and Niemann (2010) track PISA reception in one media source in each of 22 countries. They then look at the way in which PISA produced a ‘shock’ in Germany but absolutely no response in the United States.¹² Fladmoe (2011) looks at PISA reception in the media in Norway, Sweden and Finland. She finds that public television news is positively associated with PISA awareness, and public opinion is dependent, in part, on news consumption. Stack (2006; 2007) argues that Canadian and American media coverage is selective and in line with government and business interests. Takayama, Waldow and Sung (2013; 2014) draw on media accounts in three countries to

¹² Their sample is different from my own in the US case. Martens and Niemann use the search engine Factiva to search for articles about PISA in the New York Times from 2001-2008, but return zero articles from their search. I also used Factiva to search the New York Times but returned four articles from the same period.

provide a comparative analysis of the responses to PISA results in Australia, Germany and South Korea.

Dixon, Arndt, Mullers, Vakkuri, Englom-Pelkkala, and Hood (2013) compare press responses to rankings on PISA and PIRLS scores in four European countries. Their findings are mixed, as they determine that political and institutional contexts influence the tenor of the media portrayal and subsequent political responses. Additionally, Bieber and Martens (2011) and Dobbins and Martens (2012) argue that in Switzerland and France, respectively, PISA provided a new impetus in the policy process re-energizing ongoing reform efforts. Each of these studies accentuates the role of the press in PISA reception and the analyses generally agree that press coverage of PISA “produce(s) a set of conclusions, definitions of ‘good’ or ‘bad’ educational systems” (Nóvoa and Yariv-Mashal, 2003; 425). None of these reception studies offers examples of direct policy change or policy transfer as a result of coverage about PISA in the media.

Limitations to PISA Reception in the Media

These reception studies offer a contextualized approach to reception but they are somewhat limited in their data and methods. Media are expected to be an important site of PISA reception, and it is appropriate that the studies analyze the media. However this approach is also somewhat limiting. First, the research tends to look at a limited number of media outlets (sometimes only one newspaper, usually elite/high quality) and does not address the importance of a range of sources, including education blogs related to more

traditional media outlets. Additionally, the research often focuses on a specific point in time, around the release of a specific cycle of PISA scores. This limits the amount of variation in the data and therefore limits the understanding of reception. While some of the literature does make mention of additional sources, like government documents (e.g., Pons, 2012; Bieber and Martens, 2011; Dobbins and Martens, 2012; Waldow, 2009), or mentions additional methods, like interviews (e.g., Grek, 2009; Sellar and Lingard, 2013; Engel and Frizzell, 2015), it does not always capitalize on the richness and variation of these additional data.

PISA Reception as Evidence Usage

One way to reframe the focus of PISA reception studies is to consider PISA performance scores as evidence rather than governance. This approach begins by acknowledging the power of ranking and performance indicators, and the political and social influence they may have in an international context. But it emphasizes PISA data as information that may be put to use by specific actors, at specific times, for specific reasons. If PISA is understood as a new form of evidence, the questions about its influence shift. Rather than presupposing a ‘PISA effect,’ one can ask whether PISA data is being used as evidence, and if so, how? When is PISA activated? By whom? Are actors using PISA evidence similarly or in different ways? Under what circumstances? How should its use be measured? If PISA is a new form of evidence, what role, if any, does it play in education politics and policy?

On this view, PISA reception is understood as part of the shifting, complex, localized arrangements that undergird politics and policy. In focusing on PISA as evidence, rather than governance, I am emphasizing that data and numbers are powerful, but they do not govern; people do. Evidence is contextually based and selectively used. As the National Research Council (2012) points out, use of evidence is not the inevitable outcome of having evidence. They suggest that “evidence-influenced politics” may be a more informative metaphor, descriptively and prescriptively, than evidence-based policy (National Research Council, 2012, 4). The question becomes when and how PISA data is used – by actors, embedded in institutions, with particular interests, at certain points in time -- in the process of governing, and why actors find PISA a useful form of evidence.

Performance Data as Evidence

The use of performance data, in the form of indicators, has grown not just in education but across many different sectors. There are indicators for human development, human rights, governance, corruption, human trafficking and school quality, to name just a few. One recent study identified 178 such indicators, and the list is growing (Bandura, in Kelley and Simmons, 2015, 55). These data are often presented as “a named collection of rank-ordered data that purports to represent the past or projected performance of different units” (Davis, et.al, 2012, 6). As mentioned previously, the data simplify complex social phenomena so that units of analysis become comparable. Indicators are usually designed to be evaluative in that they are meant to judge performance on metrics deemed to be of importance. They embody theoretical claims about what societies value and are “markers

for larger policy ideas” (Davis, et.al., 2012, 9). While the indicators themselves are evaluative, they become especially powerful when they are presented in a ranked list, as PISA scores are. Rankings not only collapse context and meaning behind the data into one indicator, they offer a simple, visual form of comparison that encourages discussions of ‘best’ and ‘worst.’

Indicators and the Policy Process

Kelley and Simmons (2015) offer three distinct mechanisms through which ranking and indicators can affect policy. First, they can influence domestic politics. To the extent that ranking and indicators about an issue are seen as salient in the domestic arena, they can be mobilized to retain political support for policies that seem to be working, or to mobilize actors around alternative reforms. Second, ranking and indicators work through naming, blaming and shaming. Performance data can be used to identify failed policies, or the individuals responsible for those policies, triggering public opprobrium and pressure for change. In a Foucauldian sense, indicators monitor performance as a form of social control with a potential to both blame and shame those individuals and policies that do not meet expectations. This mechanism applies public and social pressure on elites to compel change. The third mechanism through which ranking and indicators can affect policy is by activating transnational pressures. The most notable example of this is credit ratings, which can dramatically influence the activities of markets. But transnational pressure may also come through third party pressure; for instance the United States uses indicators and rankings to decide whether a country should receive Millennium

Challenge Corporation funds (Kelley and Simmons, 2015, 59). Through these mechanisms, actors can use indicators to exercise power in efforts to influence policy. How indicators and ranking are used depends on the political context.

PISA and the Policy Process

PISA data is collected systematically, is simple, is quantitative, and appears objective, as does other performance data. Following Kelley and Simmons' model, it should therefore be potentially influential in domestic politics. Thus to explore the case of PISA reception in the United States, I focus specifically on the purposive use of PISA as evidence in domestic politics. In Kingdon's classic formulation, evidence is key to agenda setting, most especially at the stage of problem definition (Kingdon, 2011). Policy entrepreneurs work to align their preferred solutions to pressing problems within the larger context of politics. An alternative approach sees policy change as coming through the introduction of new ideas and upending an otherwise stable system (Baumgartner and Jones, 1993). Regardless, problem definition is a contested process based on negotiations about ideas and evidence amongst claimants. In both of these accounts, political agendas are set by those ideas and evidence that come to be seen as most powerful in defining a problem.

In this process, information is not neutral; actors use it purposefully to define problems and solutions. On this view, ideas are created, changed, and fought over in discourse. Political actors use narrative story lines as symbolic devices to manipulate issues. Ideas come in many forms, all the while making it seem as if they are simply factual descriptions of the problem in question. It is also dependent on who is making the claims.

Stone (2001) argues that all of these factors help to determine why some ideas resonate and others do not. Thus the battle of problem definition ideas is in part a battle amongst different constructions of evidence.

The second mechanism through which PISA is theorized to affect the policy process is through naming, blaming and shaming. PISA data would be used as a lever to compel policy education policy change by taking to account specific policies that are deemed to be responsible for unacceptable PISA performance. This mechanism dovetails with the externalization approach that I discuss in the next section.

The third mechanism, transnational third party pressure, is somewhat limited with regard to PISA. As I argued earlier, the OECD has little material leverage in national education systems. While PISA is marketed as a policy tool, the existing research suggests that little policy transfer or even policy inspiration has been taken from PISA thus far. The potential for this mechanism to be at play with regard to PISA seems to be somewhat limited.

Nonetheless, PISA data are potentially a very powerful form of evidence. In education, as in all spheres of public policy, quantitative data now define both problems and solutions about what is educationally legitimate.¹³ The growing numbers of international assessment instruments, the growth in numbers of countries participating in international

¹³ As Leon Weiseltier recently put it, “There are “metrics” for phenomena that cannot be metrically measured. Numerical values are assigned to things that cannot be captured by numbers. Economic concepts go rampaging through noneconomic realm...Where wisdom once was, quantification will now be. Quantification is the most overwhelming influence upon the contemporary American understanding of, well, everything.” (Weiseltier, 2015, 1, 14.)

assessments, and the calls for evidence-based policy are all in line with the belief in a rational science represented in numbers. In light of this, PISA data is well situated to be used, in Gusfield's (1981) terms, as a form of rhetoric calculated to induce belief. The process of meaning-making about PISA can be viewed as a function of defining educational problems and solutions for the political agenda.

Political Discourse and Agenda Setting

To understand how politics influences the use of PISA as evidence, and how PISA evidence influences local politics, I propose looking at political discourse. If the purpose of policy is to guide decisions and behavior, policy makers must have a notion of what kinds of decisions and behaviors are best. The deliberations about which ideas are best have typically been characterized as political discourse. One way to understand PISA reception, then, is to examine the political discourse around PISA. Following Schmidt (2008), I employ the term discourse to connote not only the content of ideas, but also the interactive processes through which ideas are conveyed. On this understanding, discourse is conceived of much more broadly than the textual interpretations of post-modernists and post-structuralists. Instead, discourse includes both substantive knowledge being transmitted and the processes by which that transmission takes place; in other words, it describes the political interaction of evidence and the actors who use it.

Externalization and Referencing

As a way of thinking about discourse, I suggest drawing on systems theory (that espouses looking closely and deeply at the local context in order to understand reception) and the concept of externalization. Comparative education researchers have drawn on the concept of externalization to guide their work in understanding reception more clearly. The externalization framework emphasizes “systemic meaning, discursive formations, semantic constructions and context specific reception” (Rappleye, 2012, 123-4). Externalization refers to the process by which actors in an education system mention situations or ideas external to education as a way to generate reform pressure.

Applied to comparative education, this approach has spawned a body of research that focuses on agency and power in local contexts. Its theoretical roots are in systems theory, as laid out by German sociologist Niklas Luhmann. Luhmann defines education as one functional sub-system of modern society. Communication binds the sub-system; when education actors communicate ideas from outside the boundaries of the education system itself, the new ideas irritate the system. This process, as elaborated by Jürgen Schriewer (1990), explains how one country borrows or lends (or purports to borrow or lend) an education policy idea or program to/from another country or institution. The process of externalization is seen as a signal that domestic political actors are looking to gain leverage in legitimizing (or de-legitimizing) domestic education reforms (Steiner-Khamsi, 2004; Waldow, 2012).

Externalization is thus a discursive action in which actors make reference to ideas or policies outside their home education system in order to generate local reform pressure. Luhmann identified several possible external points of reference that actors might use, including references to values, to science, and to institutional structures. Schriewer added externalization to “world situations,” commonly used in comparative education research (Schriewer, in Waldow, 2012, 418). Recently references to ‘international standards’ and ‘best practices’ have become increasingly common (Steiner-Khamsi, 2014). A key, albeit simplified, point regarding such referencing is that the reference itself is simply the allusion to, citation of, or comparison with something external to the education system. It does not mean that a substantive policy choice is actually being made or transferred (Waldow, 2012, 419).

Reference Societies. Theoretically, these acts of reference are believed to be important because they are made and “mobilized at particular moments of protracted policy conflict to generate reform pressure and build policy coalitions” (Steiner-Khamsi, 2014, 157). The choice of external reference reflects how actors within the education sub-system feel they can best legitimize national policy measures (Waldow, 2012). In comparative education, the use of reference societies has generally been analyzed to understand how actors legitimate their policy preferences while ‘borrowing’ policy from elsewhere. Reference societies provide a “blank canvas” (Smithers, 2004), a “projection screen,” (Waldow, 2012), or a “multiaccentual signifier” (Rappleye, 2012) that are used as evidence to justify policy preferences. While comparative education has generally

concerned itself with the borrowing and lending of educational policies and practices, less attention has been paid to the use of references in the construction of discourse (Silova, 2006 being an exception). When viewed through a discursive lens, reference societies are used strategically to help define policy problems and solutions. Reference societies are important as both reflections of particular policy preferences and as potential sources of domestic policy solutions.

Schriewer and Martinez (2004) coined the term ‘socio-logic’ to describe the contextually bound reasoning behind the choice of references in educational knowledge production. Actors, whether they are academics, policy makers, or advocates, employ differing arguments (logics) depending on the time, place, and institutional and political culture in which they are situated. Arguing that knowledge production is shaped by historical and cultural factors, and that educational research responds to the problems and conditions of a specific society at a specific point in time, their project analyzed the contents of three educational journals from each of three ‘civilizations’ (the Soviet Union, China, and Spain) over an approximately seventy-year period. They found that as political and ideological conditions shifted in those countries during the twentieth century, so too did the references in those countries’ academic literatures shift. References that in one political era were extolled as models (e.g. John Dewey in pre-Communist China) were repositioned as negative examples as political systems changed. Their results showed that over time, far from converging to a common educational ideal, journal articles

referenced educational systems in other countries for reasons and purposes directly related to the political and ideological contexts of the times

References to PISA Top Scorers. With regard to PISA, one point of externalization is references to top scorers, or countries that perform well on the assessment (Takayama, Waldow, and Sung, 2013; 2014; Sellar and Lingard, 2013). What is especially interesting is which countries reference each other's PISA performance results in their educational discourse, since an examination of the reference made reveals the underlying interests and logic of the referee. The most thorough reception studies have employed an analysis that highlights externalization to top scorers. These reception studies attempt to understand how local political, economic and cultural rationales create a logic that drives local meaning and interpretation around PISA data (Waldow, 2012).

Finland, for example, as a consistently top scorer on PISA, has become a prevalent reference society in the PISA discourse. From the time that PISA was first administered in 2000, Finland has been a top scoring country in every round of testing. A narrative of "Finnish educational success" has emerged from Finland's consistently high rank, and has taken hold in international education policy debates. Recent literature has demonstrated how the idea of "Finnish success" has played a key role in domestic policy discourse in different country contexts.

In the French case, Dobbins and Martens (2010) show how unions and political groups draw on PISA results in their rhetorical posturing while the Sarkozy government uses the Finnish case selectively to promote policy change around teacher and school autonomy. Takayama (2008; 2010; 2012) provides a detailed analysis of the Japanese case. He uses press accounts to trace the construction of a crisis discourse in Japan, chronicling how progressives in particular used Finnish success as a rhetorical strategy of protest against conservative reforms. Takayama, Waldow and Sung (2014) show that the timing and intensity with which results were covered in the national medias of Australia, Germany and Korea varied considerably. The Finnish success trope is used to construct ‘crisis’ narratives in domestic discourse that are “differently shaped by the preexisting of configuration of meanings in each country” (Takayama, Waldow and Sung, 2014, 33). The authors note two major commonalities in the cases: first, that the reference to “Finnish success” is used in the media discussions of the most contentious domestic education policy debates; and second, that all three countries show a linkage between the use of ‘crisis talk’ and foreign models in education policy debates.

Contextualizing PISA Reception in the United States

The synthetic approach I am suggesting combines the theoretical understandings of indicators with the externalization framework from comparative education, and applies this to the agenda setting stage of the policy process. Specifically I emphasize PISA data as being potentially influential in defining educational problems and solutions. However, interpreting PISA reception in the US requires an understanding of the American

educational context. In this last section of the chapter, I provide a brief overview of some of important trends since the early 1980s in American education politics and education discourse. Understanding these trends appropriately situates the PISA discourse as part of a longer-term horizon of ideas and actions that underpin the reception of PISA today.

The Education Policy Context in the US, 1983-2012

The American reception of PISA data occurs in the context of long-held beliefs and ideas about American education. To understand PISA reception in the United States, it is necessary contextualize the American education policy landscape during the PISA era. To that end, I discuss four trends that help situate my analysis of the PISA narrative in the United States: the use of comparison as a strategy to problematize American education; a focus on assessment as a form of accountability; the tempering of traditional partisan positions in education politics; and shifts in the locus of activity around education policy.

Comparison as Problematization

Comparison has a potent history in the United States of shaping and problematizing education. It was a major preoccupation of education experts in the early and mid-19th century (Noah, 1985) and continued through the twentieth, with reports from travellers such as Horace Mann through John Dewey and others. American education thinkers have long been interested in how education is done in other places, and what they might

learn from those systems.¹⁴ In more recent times, education rhetoric has invoked comparison to highlight education as key to American national interests. The comparisons and rhetoric have changed over time, reflecting the cultural and political overtones of the era. However I focus here on the ‘crisis’ of US education that began in 1983¹⁵ with the release of *A Nation at Risk (NAR)*.

Crisis rhetoric in US education. The tumultuous decade of the 1970s saw the US coming out of the Vietnam War, into the oil crisis, and disillusioned with the level of success of the Great Society programs of the 1960s (Vinovskis, 2009). Additionally, the US economy was in a recession; inflation and unemployment were high, productivity was down, and real income declined. American economic woes were only heightened as the press showered attention on the growing Japanese economy, which was avoiding double-digit inflation and growing their semiconductor and electronics industries to boot.

Against this backdrop of fear and uncertainty both domestically and internationally, US Secretary of Education Terrell Bell commissioned a blue-ribbon panel of business and education leaders to issue a report on the state of American education. The resulting

¹⁴ In fact, a whole field of study, that of comparative education, developed to address the very issue of how we can know and what we can learn from other systems of education.

¹⁵ A discussion of the role of crisis talk in American education could easily begin before the 1980s. Most obviously in the 20th century, the Sputnik crisis of the 1950s not only involved hysterical rhetoric, but also directly impacted the US education system. When the Soviet satellite was launched, many saw it as reflecting the failure of the US education system, especially in math and science. Subsequent analysis of the American system focused on academic excellence, and concluded that the US was far behind the Soviet Union (see Trace, 1962). The Sputnik crisis framed education in terms of national security concerns, which “fueled a political movement to beef up the academic performance” (Chubb and Moe, 1990, 7). Sputnik also led to the rapid passage of the National Defense Act of 1958, which included increased educational funding and increased education policy influence for the federal government.

publication, *A Nation at Risk: The Imperative for Educational Reform* (National Commission on Excellence in Education, 1983), would prove seminal in framing the understandings of the US education system that remain today. The report painted an alarmist picture of America's education system, famously decrying "a rising tide of mediocrity that threatens our very future as a Nation and a people." It went on, "What was unimaginable a generation ago has begun to occur – others are matching and surpassing our educational attainments" (National Commission on Excellence in Education, 1983, 4).

As scholars have shown, *NAR* led to an enormous shift in American education politics because it redefined the agenda status of education (Mehta, 2006; Vinovskis; 2009; Guthrie and Springer, 2004; Manna, 2006). Released during a recession, it offered an explanation of the relative success of such international competitors as Japan¹⁶ and Germany. Underperformance in education was linked directly to the safety and security of the country vis-à-vis economic growth. In framing US education in an international comparative perspective, the economic goals of schooling were put front and center and were embraced by the business-oriented right. The focus on excellence for all students overtook debates about poverty and race that had dominated the previous decades (Mehta, 2006).

¹⁶ As Williamson recounts, in the 1970s US productivity was down to 1.9% annually from its long-term historical average of about 2.3%. At the same time, Japan experienced a 5% growth increase. Thus a narrative of Japanese strength and American decline was borne. This was in spite of the fact that Japanese growth was down from its growth rate in the 1960s, which had been about 10% annually (Williamson, in Feuer, 2011, 15).

Domestically, the report changed the agenda status of education, making it a national agenda item.¹⁷ Due in large part to the decentralized, federal system of government that leaves educational decision making powers in the hands of local and state officials, the US federal government has historically played a limited role in education policy. With the headlines from *NAR*, education policy became a cause for policy makers at every level, and US presidents now prided themselves on being “education presidents.” *NAR* also brought together previously disparate groups – legislators and business leaders – around a new common cause and a new language of education. This increased status of education prefaced greater federal involvement in what had been a highly decentralized system. In so doing, it set up a discourse that insisted on academic achievement as a bulwark against the pressures of an increasingly globalized and economically competitive world.

NAR was a media sensation. While some claimed that the crisis it invoked was artificial (Berliner and Biddle, 1996), its narrative became the dominant trope in American education discourse. The report put its findings of American decline in a rhetorical context of international underperformance. American schools were failing their students and, as a result, America’s economic competitiveness was suffering. Building on the fear and anxiety of the era, the report specifically referenced Japan, South Korea and Germany as our “determined, well-educated, and strongly motivated” economic

¹⁷ Manna (2006) and McGuinn (2006) provide detailed if slightly different accounts of this process since the release of *A Nation at Risk*. Manna argues that the rise of education on the political agenda is due to a push from states to pressure the federal government to embrace reforms like standards and accountability that were fomenting at the state level. McGuinn argues that the rise in agenda status came from electoral politics, where education became a part of the vote-getting strategy of elected officials.

competition. The report also foreshadows American demands for new ways to measure and compare educational performance, both domestically and internationally.

Comparison since *A Nation at Risk*. *NAR* was enormously important for the agenda status of education and for the way in which it brought new actors together in the education policy arena. But it also marked a turning point in the way in which education was problematized. Putting American schools in international perspective heightened the sense that school improvement was not only necessary, it was possible (Mehta, 2006). Indeed the diverse members of the committee that wrote the report came together around the idea that education was the key to international economic competitiveness, and that education would provide returns to both macrolevel and individual economic growth. In the midst of a recession while Japan was thriving, *NAR* concentrated attention on the economic functions of schooling.

The linking of education to economic well-being has remained a critical part of the US education narrative and feeds a vision of comparison and competitiveness. That education fuels both macrolevel and individual success is such a common trope in education discourse it is taken for granted. Education discourse often includes metaphors of the Olympics, horse races, or league tables that directly link competition to schools. Comparative international assessments fit in to this trend easily, as increasingly sophisticated and available data makes international comparisons more common. This narrative – that America is in decline but that education provides the path to a stronger, brighter future – remains a critical part of the American educational psyche.

Accountability and Assessment

The power of comparison has been bolstered by the domestic accountability and assessment regimes that have developed since the release of *NAR*. Koretz (2008) provides a sweeping account of the role that testing has come to play in American education, as a source of performance data and as part of a larger accountability regime. As he describes, the common approach of low-stakes, diagnostic achievement testing began to change in the United States in the 1960s. With the passage of the Elementary and Secondary Education Act (ESEA) in 1965, the federal government passed far-reaching legislation authorizing funding for k-12 education throughout the US. As part of its Title I program, designed to equalize funding between those areas with economically disadvantaged students and those without, schools were mandated to evaluate students based on scores of standardized, norm-referenced achievement tests. At the same time, in the 1960s, the federal government established the National Assessment of Educational Progress (NAEP), which would eventually become to be known as “the nation’s report card.” Though NAEP was designed to provide a description of achievement trends, with no consequences for teachers, in retrospect these two programs marked a change in federal action on assessment and accountability. As Koretz notes the assessments “signified the beginning of a fundamental shift in the goals of testing, from diagnosis and local evaluation to large scale monitoring of performance and ultimately to test-based accountability” (Koretz, 2008).

As the status of education rose on the political agenda, politicians and business leaders advocated assessment as a way forward in education reform. In a paradigm where education is the basis for economic success, schools rather than social forces need to be measured and held accountable for academic performance. If schools are to be held accountable, however, “someone has to decide what level of performance is good enough” (Koretz, 2008, 69). It was not surprising then, that when the ESEA was reauthorized in 2001 as the No Child Left Behind (NCLB) act, the law required data-driven accountability measures for all schools (see Manna, 2011; DeBray, 2006; Vinovskis, 2009). One can safely say that a culture of standardized assessment now pervades American education. The prominence of a testing and accountability culture in the US dovetails with US development and support of an ILSA regime.

The Tempering of Partisanship

The decades after the publication of *NAR* also saw a collapse of traditional partisan positions in US education politics, as *NAR* redefined the parameters of what was “sensible centrist reform” (Mehta, 2013, 312). As such, partisan interests converged. Democrats, who had traditionally made education arguments around equity concerns, accepted the accountability, testing and standards movements as a way to ensure that poor and minority students were being served as well as their better-off counterparts. Republicans, who were fighting for the closure of the Department of Education in the early 1980s before *NAR*, now embraced accountability, testing and standards as talisman against threats to American competitiveness. Civil rights groups thought the new

accountability regime would help close the achievement gap; business leaders supported it in the name of increased productivity. Ideologically driven think tanks began to cross-pollinate staff and ideas (DeBray-Pelot and McGuinn, 2009). A new bi-partisan consensus emerged around a data-driven accountability and assessment agenda. Embodied in the passage of NCLB in 2001, conservatives dropped their objections to federal influence over schools and liberals supported a standards and accountability agenda as a new solution to their equity concerns.

Shifts in the Locus of Policy Activity

One consequence of the rise in agenda status of education has been the expanded role of the federal government in education politics. As McGuinn (2006) and others have noted, the federal government has historically had very little role in educational governance in the US, where the control of education has typically been left to states and localities. For instance, even Title I funds, designed to improve equity through federal funding as part of the ESEA, spread federal dollars with little accountability, so most decisions remained local (DeBray-Pelot and McGuinn, 2009).

The passage of NCLB marked a new high-water mark of involvement of the federal government. As McGuinn (2005) has written, the breadth and depth of the federal involvement in schools is a remarkable development, especially given the history of decentralized school governance in the United States. The original federal role for education as put forth by the ESEA was narrowly construed to target poor students but

was significantly transformed with NCLB. McGuinn also notes that federalism has long acted as a restraining influence on the role of the federal government in education, but now has been overtaken by the new politics of federal involvement in education.

One potential consequence of the increased involvement of the federal government in education has to do with the role of educational knowledge in national politics.

According to Henig, these structural changes have made the political environment more polarized, with perhaps an increased demand for research that “maps directly onto... core debates... if not with regards to research design and analysis then with regard to rhetoric” (Henig, 2008a, 360). This suggests that there may be an increasingly important role for evidence to support the more highly polarized institutional (rather than partisan) cleavages that define the educational landscape.

The Development of the Common Core State Standards and use of PISA

The Common Core State Standards (CCSS) lend themselves to the theoretical exploration of how PISA scores might be tied to domestic education politics. On the view of externalization, the adoption and subsequent controversial implementation of the CCSS would be the kind of contentious policy that might encourage references to PISA. Actors looking to legitimate the CCSS would make reference to PISA and to top scorers who also had national standards as a rhetorical mechanism to push forward their reform agenda.

The development of the CCSS. With NCLB's requirement of annual standardized achievement scores for all students between the third and eighth grades, the legislation seemed to resolve long-standing debates in education in favor of individual achievement (Mehta, 2006). Following a turn towards standards in 1989, after US governors convened in Charlottesville, Virginia for what would become known as the education summit, national goals and standards were on the agenda of both the states and the federal government. By 2009, those goals were drafted with the support of the Council of Chief State Schools Officers (CCSSO) and the National Governor's Association. After a brief public comment period in the spring of 2010, the Standards were released in June 2010, and immediately 40 states signed on to adopt them, incentivized by the federal grant competition known as Race to the Top (RttT). States that agreed to adopt the CCSS gained more points on their RttT applications, increasing their chances of receiving a share of the money. A total of 46 states had signed on to the initiative by 2011.

More recently, as implementation of the Common Core has progressed, backlash to the standards has grown.¹⁸ Conservatives see the initiative as an overreach of the federal role in education. Liberals see it as an overreach of an accountability regime that focuses on testing to the detriment of student learning. The largest American teachers union, initially supportive of the standards, has called for a pause in implementation. At this

¹⁸ Partisanship in opinion has grown as well. See <http://www.brookings.edu/research/papers/2015/04/16-chalkboard-common-core-west>

writing, three states have dropped out of the initiative completely, while the governors of at least four more states have publicly questioned their states' participation. And there has been a groundswell of parent and teacher resistance to the implementation of the standards and especially to the new standardized tests that align to them.

The Common Core and PISA. According to the documents discussing the criteria by which the CCSS were developed, “the standards will be informed by the content, rigor, and organization of standards of high-performing countries so that all students are prepared for succeeding in our global economy and society” (Common Core Standards Initiative, n.d.). Specifically, with regard to the role of international benchmarking and the development of the CCSS:

“International benchmarking refers to analyzing high-performing education systems and identifying ways to improve our own system based on those findings. One of the ways to analyze education systems is to compare international assessments, particularly the Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS). Prior to the development of the Common Core State Standards, research revealed striking similarities among the standards in top-performing nations, along with stark differences between those world-class expectations and the standards adopted by most U.S. states. As a result, standards from top-performing countries were consulted during the development of the Common Core State Standards. The

college- and career-ready standards appendix lists the evidence consulted” (Core Standards, 2015)

Even before the standards were released to the public, the CCSS were tied to international assessments. There is thus an acknowledgement, at least behind the scenes, that ILSAs are playing a role in domestic education, in this case providing guidance in the development of the CCSS.

Conclusion

The literature to date about PISA shows the rapid growth in PISA participation as part of a larger growth and participation in international assessments. Many argue that PISA is part of an educational ‘governance by numbers’ regime. A growing body of literature looks at PISA reception in specific national contexts, and has begun to look closely at the types of references that are made to top scoring educational systems outside of the home country.

In this chapter I am suggesting that the way to understand PISA reception in the United States is to analyze the use of PISA evidence in the context of domestic education politics. I suggest that using an externalization approach to examine domestic policy discourse helps to further theorize about PISA reception. I propose an analysis of educational discourse as a way forward in understanding how actors and ideas interact to define political problems. Of particular interest are whether and how PISA data is used

to problematize American education, and where and how the discourse then seeks political solutions to the problems it has defined.

Chapter Three: Using Content Analysis of Educational Discourse to Understand PISA Reception in the US

As I discussed in the previous chapter, a combination of theoretical approaches can be used to better understand the way in which PISA is received in a local context. I provided specifics about the American education policy landscape since the early 1980s in order to situate the trends in which PISA reception has been entrenched in the 2000s. I also suggested that American educational discourse, broadly construed, is an appropriate arena for understanding PISA reception in the United States. In this chapter, I elaborate on my model of discourse in the American education space. I concentrate on universities (the academy), think tanks and the media as creators of educational knowledge that can be used in education discourse and policy. In this chapter, I present my research design, including data collection and analytical methods. I also discuss research limitations.

Research Design

My research seeks to understand how PISA has been received and understood in the United States from 2000 through 2012. Though some basic statistical analysis is presented, I rely primarily on a qualitative research design. According to Miles and Huberman (1998) and Creswell (2007) qualitative methods provide advantages for an analysis like the one undertaken here. Qualitative methods are characterized by groundedness, where phenomena are studied within the unique political and historical contexts in which they are embedded. They are also methods that encourage understanding complexity and detail, which are appropriate as I aim to contextualize the

reception of PISA while taking into account the complexity of the American education space. Qualitative research is also appropriate in that it emphasizes in-depth knowledge across relatively few cases (Ragin, 1987), as I do in this study.

While there is an obvious tension between aiming for generalizability of results and exploring one case in-depth, I follow existing reception studies in conducting a small-N study by examining the case of the United States. In case study research the analytic approach centers on building an in-depth understanding of context and then exploring a bounded system over time through detailed analysis (Cresswell, 2007). This is a natural approach for a reception study, which seeks precisely to undertake a detailed analysis of local context. As mentioned previously, the US is chosen because of its historical background as a strong proponent of international assessment and PISA in particular, as well as its historically mediocre performance on these assessments. There has also been extremely limited research done about the US with regard to PISA reception, making it a reasonable selection for case study.

In this dissertation I take a broader view of where and how knowledge about PISA is developed. Drawing on Schmidt's (2008) model of discursive institutionalism, I argue that how PISA is interpreted locally depends not only on the media but on other actors as well, who communicate in an effort to persuade others of their views. On this view, discourse is a useful site of study because it shows how agents conceive of, and deliberate about, political phenomena. This approach sees discourse as an interactive process

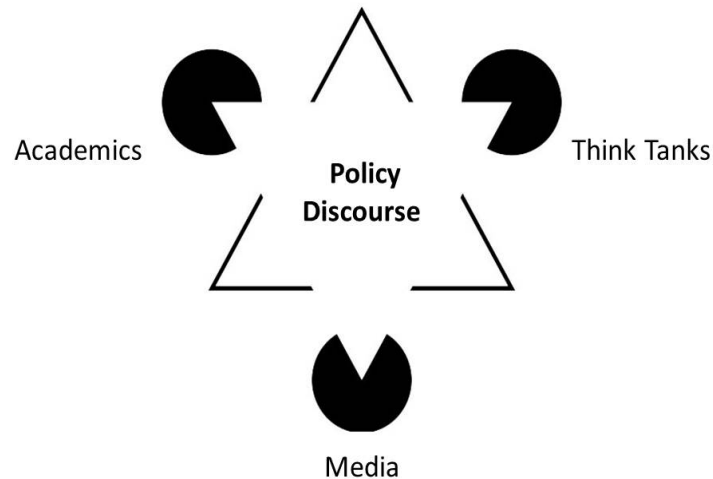
between content, context, structure and agency. It aims to understand how and why public problems are labeled as such, and how certain discourse and solutions become legitimated based on particular problem definitions.¹⁹ In the next section I provide a model of knowledge production that looks at three sets of actors, institutions and the literature they produce. Taken together, they construct an American education policy space within which understanding about PISA is created and contested.

American Policy Knowledge

In this dissertation I suggest that analyzing discourse is an appropriate way to understand PISA reception in the United States. However, previous work has taken a somewhat restrictive view on where discourse happens. Existing assessment studies focus almost exclusively on the media as the site of reception. Figure 3.1 represents a broader conception of the construction of educational policy knowledge in the form of a Kanisza triangle. This representation not only suggests the interplay of academic and think tank literature as well as the media in generating knowledge used for policy and discourse, but also hints at the constructed nature of the discourse. If one looks closely, she sees that the triangle is actually an optical illusion; there are no triangles in the picture. In using this form I mean to suggest that the ways in which PISA is understood are constructed within the preexisting norms and interests associated with each set of actors and institutions.

¹⁹ See also Gusfield (1981) and Stone (2001) for classic expositions about the formulation of public problems through rhetoric that defines causes, problems, responsibilities, and solutions.

Figure 3.1: Policy knowledge production in the American education space



Of course, the academy, think tanks and the media are not the only sites in which education discourse is constructed. Government agencies, policy analysts, members of the US Congress, unions, publishers, and private foundations, to name a few, are also involved in creating educational discourse and policy. Research has shown that links between the academy, think tanks, and the media are important in public discourse (Yettick, 2009), suggesting that studying this combination of actors and institutions in educational discourse can help produce a deep picture of the PISA narrative develops (Figure 3.2). This approach is also in line with previous knowledge utilization research that shows that the media, though important, are not the only source of policy knowledge (e.g., Rich, 2004; Levin, 2004).

Table 3.2: American education policy knowledge production

	University-based academics	Policy analysts at think tanks	Journalists and media
Products	<i>Journal articles</i>	<i>Research reports</i>	<i>Press accounts</i>
Interests in producing research	<i>Add to scholarly debate; attain professional recognition and job tenure</i>	<i>Create credibility to gain access and influence policy process</i>	<i>Report information to public; act as watchdog; sell publications</i>
Publication venues	<i>Academic journals</i>	<i>Usually self-published</i>	<i>Newspapers, magazines</i>
Quality control	<i>Peer review</i>	<i>Board review; colleague feedback</i>	<i>Editorial oversight; letters to the editor</i>
Funding source	<i>Salary</i>	<i>Private foundations, individual donors, political parties, corporations</i>	<i>Publication sales</i>
Scope of research	<i>Focused on one area of expertise</i>	<i>Focused on one, several or many areas of expertise</i>	<i>Focused broadly, though individual journalists may focus on one beat</i>

Academics and peer-reviewed literature

The effects of scholarship on education policy have been decidedly mixed. Historically, academics were seen as being neutral and above the fray of politics. In the United States, the Progressive Era reformers of the early twentieth century believed that properly trained university experts would be able to develop “real” solutions to the social and economic problems of the day, above the nastiness of patronage politics (Mehta, 2011; Hess, 2008). For much of the twentieth century, experts offered ideas and data that were thought to be rigorously crafted and credible. More recent knowledge utilization research has focused on the fact that research is always conducted in a specific organizational and political context (Hess, 2008). Though academic research may be “spun” by pundits and policy makers to support a desired view (Henig, 2008b; Whitty, 2012; Goldhaber and Brewer,

2008), the overall effects of research are ambiguous. Nonetheless, as a source of elite knowledge production, academics can be considered influential experts or symbolic analysts. In the context of international rankings, the normalizing discourse of international educational assessment is supported by academics that hold the ‘power of interpretation’ (Carney, Rappleve and Silova, 2012, 384).

Though individual academics produce the bulk of educational research, they are usually nested within the organizational structure of a university where tenure provides the primary incentive system. This system requires the publication of scholarly work in peer-reviewed journals if academics are to be granted tenure. Thus the incentive first and foremost is to produce scholarly work. Academics may be less interested in producing research that is relevant to policy makers than producing research that gets published, by debunking a common theoretical stance or focusing on methodological issues, for instance.

Policy analysts and think tank literature

Think tanks may be defined as “independent, non-profit organizations that produce...expertise and ideas to support and influence the policymaking process” (Rich, 2004, 12). The rise of think tanks in the United States began in the Progressive Era of the early twentieth century, when they were formed as a source of objective, rational, and politically neutral expertise to solve the pressing problems of the post-Industrial Revolution. According to Rich (2004), during the thirty-year span between 1970 and

2000 the number of think tanks quadrupled from fewer than 70 to more than 300. During the same period the role of the think tank analyst shifted from being exclusively a detached analyst to frequently becoming an aggressive political advocate. As Rich notes, think tanks are concerned with promoting values. Those think tanks that are more value-laden -- typically conservative think tanks -- have received more traction in their ideas and proposals than those think tanks -- typically progressive -- that feel they need to be value neutral and objective.

Think tanks are incentivized to produce high quality research in order to build reputation and gain influence amongst policy makers, other analysts, journalists, funders and other influential elites. Marketing and dissemination, with an eye towards the timely release of research on 'hot' topics, is of crucial importance. Unlike the publication process in academic journals, think tanks often draw on the speed and reach of electronic media for rapid and widespread dissemination of their research. Individual researchers may also be quite actively engaged in political debate. A recent study found that while universities produce 14 to 16 times more educational research than think tanks, think tank research was substantially more likely to be mentioned in the press (Yettick, 2009). Once the think tank research has been recognized, either by decision makers or in the popular press, the think tank gains greater access to policy makers, which in turn increases the think tank's influence.

Journalists and the media

As I have discussed, media play a key role in interpreting and disseminating ideas about public policy (Henig, 2008b; Snyder and Ballentine, 1996; Ferree, Gamson, Gerhards, and Rucht, 2002). Much has been written about the role of the media in shaping public opinion, setting agendas and informing public policy decisions. Most reception studies have understandably relied almost exclusively on the media as the site of reception. Research has long demonstrated the key role that media play in political agenda-setting by choosing which stories and issues are reported on, and then how those issues are covered (see McCombs and Shaw, 1974; Gamson and Modigliani, 1989; Gamson, 1992). This coverage both reflects and creates public culture and public opinion. Additional work has shown the role media play in framing (defining and interpreting problems and solutions; Benford and Snow, 2000) and priming (how the agenda set by media affects public opinion; Chong and Druckman, 2007). Some have argued that mass media are the primary actors in the US public sphere for providing political interpretation (Ferree, et al., 2002, 63).

Sampling

In order to examine the reception of PISA in US education discourse, I conducted a content analysis on texts in the American education policy space published from 2000 to 2012. This time frame coincides with the first decade of PISA assessment during which four cycles of PISA were administered (2000, 2003, 2006 and 2009) and their results published. To conduct the analysis I created an original dataset of texts from the three

aforementioned bodies of literature: academic education journals, think tank publications and American newspapers.

Two things distinguish my sampling methods from previous work. Across all sources, I sample all literature published between 2000 and 2012. This twelve-year period corresponds to the release of results on the first four rounds of PISA.²⁰ By sampling across the first decade of PISA, I am able to provide both a more nuanced and more systematic approach to understanding reception. As I demonstrate, time is an important variable in understanding reception.

Additionally, I employ a broader definition of what a reference to PISA is. Extant literature on PISA reception is not always methodologically clear, but the results suggest that data is collected only from those articles that offer an in-depth discussion of PISA, meaning that the text is focused exclusively on PISA and PISA results.²¹ I believe this approach is too restrictive. It is my contention that meaning-making about PISA is constructed not only from in-depth discussions of PISA but also in simpler, less comprehensive references. These simple mentions provide a subtler look at how ideas

²⁰ PISA is administered triennially. However PISA results are released one year after the date of administration (e.g., PISA 2000 was administered in December 2000, with the results released in December 2001). This means that my sample, which runs through 2012, covers discourse up through the 2009 cycle. PISA 2012 results were not released until 2013, a date beyond my sampling parameters.

²¹ For example, Takayama (2010) specifically drops from his sample “references to Finnish education reform [that] were not substantive enough for my discourse analysis” (p. 54). This results in a newspaper sample of 24 articles for a seven-year period. On the US case, Martens and Niemann (2010) report finding no articles from the New York Times on PISA from 2001-2008. Sellar and Lingard (2013a) also take a restricted approach, analyzing one New York Times article, one press release and the 2011 State of the Union address to gauge US response to the 2009 PISA cycle results.

about PISA travel and are created. By broadening the definition of a reference I am able to get at this more nuanced understanding as well as gather a more robust sample.²² I detail the selection process for each data source, by literature type, below.

American academic literature

The first sets of texts I analyze are peer-reviewed journal articles from three leading American education policy journals. While there is no clear agreement on the flagship journals in the US education field (Henig, 2008a), I choose the American Educational Research Association (AERA) as an association whose journals can reasonably represent educational knowledge production in the US.²³ The AERA is the professional association of education researchers in the United States. It has more than 25,000 members, including faculty, researchers, graduate students and other education professionals and is the major national research society in American education. According to its website, the AERA “strives to advance knowledge about education, to encourage scholarly inquiry related to education, and to promote the use of research to improve education and serve the public good” (AERA, 2015).

²² This approach is also different from the very limited literature that analyzes academic journal articles. For instance, Owens (2011) analyzes PISA in scholarly journal articles, but only when PISA data is used as a source for secondary analysis. Lockheed (2013) also analyzes scholarly references and articles to PISA in academic journals but compares proportions of citations (n-grams) referring to PISA and TIMSS.

²³ A journal’s impact factor is a common way to determine the importance of a given journal. The ranking calculates the average number of citations per article in the previous two years. Though its accuracy has come under fire for the fact many journals have figured out how to artificially inflate their journal impact factor by publishing certain types of articles, it is often used to assess importance and impact. I do not employ it here, however, as I wanted to bound the universe of education journals to the American education space through the AERA. As a much smaller field, the choice of journals in comparative education was easier as I chose the two most prominent journals in the field (though because of the field’s size, their impact factors are rather low).

To be sure, AERA represents only one part of the academy. The academy as a site of knowledge production is a big place. Education researchers are interested in different questions than economists, for instance, and bring certain orientations and biases to their research about PISA. AERA also represents an elite group by its very nature, those academics that have an advanced degree and produce scholarly research. These limits notwithstanding, AERA covers the research and interest of thousands of academics in the United States or publishing in American publications. This makes it an appropriate choice from which to draw texts from the American educational space.

The AERA publishes five peer-reviewed journals a year (Table 3.3). I sample three journals that publish articles on educational policy: the American Education Research journal (AERJ), Educational Researcher (ER), and Educational Evaluation and Policy Analysis (EEPA). I chose these journals because they each devote some focus of their publications to education policy. The other journals published by AERA are not as strongly focused on policy research and are therefore excluded.

To identify relevant articles, I searched the publisher's database, Sage Premier 2010, for each journal, running an Advanced Search using the terms "Programme for International Student Assessment" or "PISA," and "OECD PISA" in "all fields." The time period was delimited from January 1, 2000 through December 31, 2012.

Table 3.3: Overview of AERA journal sample

<i>AERA Journal Operations, 2010</i>								
Journal	Manuscripts			Published Articles			Subscriptions	
	No. Received	No. Accepted	% Accepted ^a	No. Rejected	Items Published	Total Pages Published	Member and Nonmember	Institutional ^b
<i>American Educational Research Journal</i>					27	968	11,520	6,840
SIA section	196	16	8.2	162	10	321		
TLHD section	207	16	7.7	93	17	647		
<i>Educational Evaluation and Policy Analysis</i>	212	22	10.4	102	23	520	3,899	6,414
<i>Educational Researcher</i>	220	38	17.3	142	53 ^c	706	24,610	6,593
<i>Journal of Educational and Behavioral Statistics</i>	114	65	57.0	67	36	774	2,061	6,314
<i>Review of Educational Research</i>	189	12	6.3	68	18	638	7,538	6,948

Source: AERA, 2011.

a= Manuscript counts and rates are within one calendar year.

b=Includes discounted and free institutional subscriptions.

In order to compare American academic literature with the broader literature in comparative education, I sampled texts from comparative education journals to assess whether and how those journals, for which large-scale international assessments are a regular topic of interest, publish articles about PISA. I include a sampling of articles from Comparative Education Review (CER), the flagship journal of the Comparative and International Education Society, and Comparative Education (CE), another leading comparative education journal. These journals publish both American and international authors, and provide an interesting counterpoint to the literature generated with only an American focus. The searches in these journals employed the same search methods and parameters as in the American academic journal articles.

Think tank literature

I collected a sample of texts published between 2000-2012 by two American education think tanks. Non-peer reviewed policy analysis literature plays an increasingly important role in disseminating ideas and influencing policymakers (Rich, 2004; Rothstein and Hopewell, 2009) and as such is an appropriate source for my data. For the purposes of this dissertation, I define “policy-oriented think tank” as a non-profit organization engaged in public policy research on education. Initially, I performed a Google search for public policy and education policy think tanks. I then visited the websites of each of the top ten results, assessing their mission statements, their work in education, and their political orientation (as described in the popular press or on their websites) and whether or not they publish research. I also cross-referenced my search results with Swanson and Barlage (2006), Yettick (2009) and McGann (2013), each of which includes rankings of influential education policy think tanks. I then chose two prominent think tanks from which to sample text: the Brookings Institution and the Thomas B. Fordham Institute, each of which publishes education research, is recognized as influential in the literature, and represents different points on the political spectrum (Table 3.4).²⁴

The Brookings Institution (hereafter, Brookings) was founded in 1916 as the “first private organization devoted to analyzing public policy issues at the national level” (Brookings

²⁴ While the partisan distinctions between Brookings and Fordham are supported by the descriptions on their websites and are triangulated with Swanson and Barlage (2006) and Yettick (2011), in fact their positions on education during the past decade may have been closer than the partisan descriptions suggest.

Institution, 2015). Brookings is a private, nonprofit organization “devoted to independent research and innovative policy solutions” (Brookings, 2015). Brookings is generally considered a center-left leaning organization that is consistently ranked as one of the most prestigious US think tanks (McGann, 2013). Brookings produces education research in its topical education program, and also through the Brown Center on Education Policy.

Table 3.4: Overview of think tank sample

Think Tank	Location	Mission	Political Leaning
The Brookings Institute	Washington, DC	“The economic and political well-being of any society requires a well-educated citizenry. Brookings’s work has extended beyond the K-12 bookends to include preschool interventions, higher education and the challenges of education in developing countries. Experts are tackling fundamental issues on the role of education in the national and global economy.”*	Center-left, liberal-leaning
The Thomas B. Fordham Institute	Washington, DC/Dayton, Ohio	“The Thomas B. Fordham Institute is the nation's leader in advancing educational excellence for every child through quality research, analysis, and commentary, as well as on-the-ground action and advocacy in Ohio.”**	Right-leaning, Conservative

* See <http://www.brookings.edu/research/topics/education>

** See <http://www.edexcellence.net/fordham-mission>

The mission of the Brown Center is to “bring rigorous empirical analysis to bear on education policy in the United States.” It conducts wide-ranging educational research on timely topics that are of particular interest to education policy makers (Brookings, 2015). Among other research products, the Brown Center annually produces a Report on

American Education, which has focused in part on American performance on international assessments during the past several years.²⁵ Brookings is funded through a combination of foundation, corporate, individual and government donations.

The Thomas B. Fordham Institute (hereafter, Fordham) was founded to advance “educational excellence for every child through quality research, analysis, and commentary” and seeks to promote education by “producing rigorous policy research and incisive analysis” and “building coalitions with policy makers, donors and organizations and others who share our vision” (The Thomas B. Fordham Institute, 2015). Fordham is considered a right-leaning, conservative think tank (Yettick, 2009) and is considered a leading organization on educational policy. Fordham is supported through grants from public and private foundations as well as individual donors.

Similar to the academic journal article search, I searched the website of each think tank using the terms “Programme for International Student Assessment” or “PISA,” “OECD PISA” in “all fields.” As in the academic journal article search, including the terms “US” or “United States” did not increase the search results and so were excluded. The time period was delimited from January 1, 2000 through December 31, 2012.

One important limitation to the think tank data collection process had to do with the life cycle of Internet archives. I collected think tank data in 2012-2013. It is quite possible

²⁵ The author of these reports is Tom Loveless, who was previously the US representative to the IEA, which produces TIMSS and other non-PISA international assessments.

that materials from the early 2000s were removed from the think tanks' web servers and were unavailable in their archives. If this is the case, it would introduce selection bias into the think tank data.

American media

For the sample of media articles in the US, I choose four leading US newspapers, all dailies, and each considered a paper of record, from across the political spectrum (see Table 3.5). *The New York Times* is considered an elite liberal newspaper; *The Wall Street Journal* is considered elite conservative. Additionally, *The Washington Post* is included because of its elite coverage of politics and policy due to its location in the nation's capital. These sources are known for their influence and are often used in studies of American discourse (e.g., Fiss and Hirsch, 2005). Finally, *USA Today* is considered a more popular press with short, dramatic news coverage.

Table 3.5: Overview of media sample

Source	Frequency of publication	Political orientation*	Circulation** (includes both print and digital circulation)
New York Times	Daily	Left/liberal	1,865,318
Wall Street Journal	Daily	Right/conservative	2, 378,827
Washington Post	Daily	Left of center	473,462
USA Today	Daily	Centrist	1,674,306

*Political orientation of media coverage can be contested. My categorizations are drawn from a variety of sources, including Swanson and Barlage (2006), McGann (2013) and Yettick (2009).

**Circulation figures from Alliance for Audited Media, 2013.

Using Factiva, a subscription-based search engine owned by the Dow Jones Company, I was able to search all four newspapers' databases. As with the academic and think tank

texts, I searched for “PISA,” “Programme for International Student Assessment” and “OECD PISA.” Unsure that Factiva was actually capturing all articles and blog posts from each media source, I searched the online archives of each media source manually, using the same search terms. This uncovered several more articles that included the search terms but had not been retrieved in the initial searches. As with the academic and think tank literature, the time period was delimited from January 1, 2000 through December 31, 2012.

After each search was completed, results were reviewed and catalogued. Abstracts and/or introductions were read to determine whether or not the article should be included in the sample. If I was unable to determine whether or not the article/publication was appropriate, I proceeded to read the entire article to make that determination.

Exclusion rules

I eliminated results outside the specified date range. Additionally, I excluded results that were duplicates or were no longer retrievable from the server. When there was a question as to whether a text was appropriate for inclusion in the sample, I employed a series of decision rules. These included the exclusion of search results that were book reviews, errata, advertisements, conference invitations or announcements. Results were excluded when the reference to PISA appeared only in a footnote, or only as part of a graph showing results without any explanation or interpretation. Any results from transcripts of spoken (testimony, live events) text rather than written texts were excluded.

When articles appeared in two different formats (meaning in print as well as online) the article was counted once but the different formats were noted and coded as such.

Additionally, blog posts were included from both the think tanks and newspapers as an important and timely source of data and information on PISA discourse.

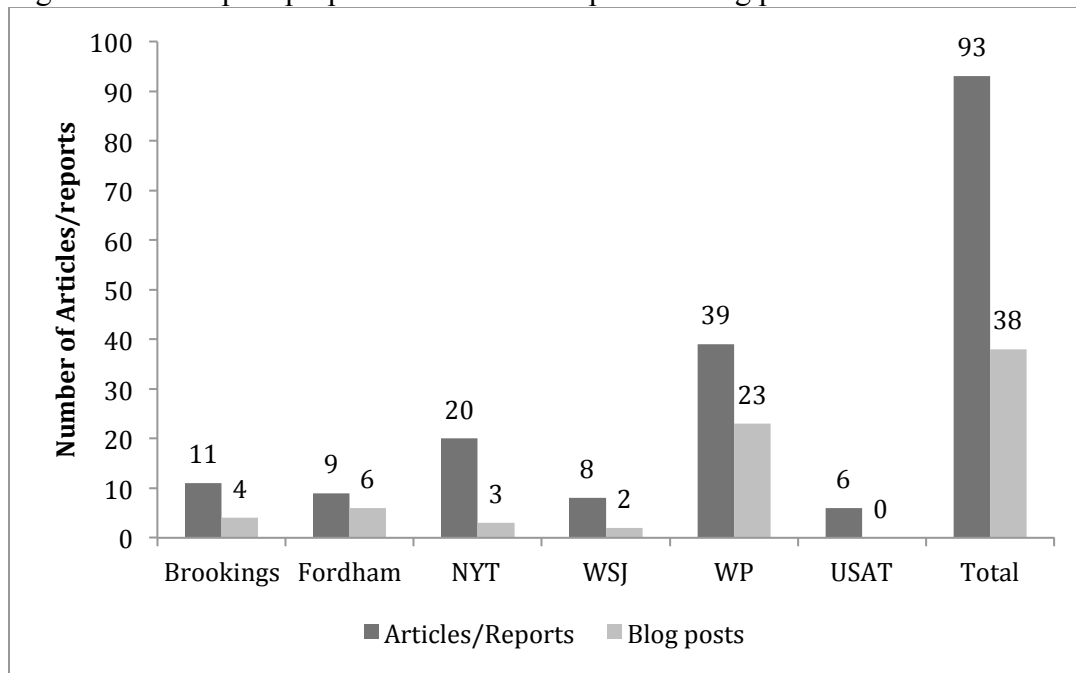
Overall sample

The sampling resulted in a total of 130 texts to be coded for analysis. About half of the texts are from media sources (N=73), while approximately thirty-one percent are from the academic literature (N=37), and approximately seventeen percent are from the think tank literature (N=20). While the largest number of texts comes from the media, this likely represents an artifact of the media publication cycle. The daily publication cycle of media sources means that there are many more media articles than think tank reports and academic journal articles available to search within the prescribed date range. Academic journals have a publication cycles that run in months, or even years, rather than days. Think tanks have more leeway, but significant reports require some amount of time to ready for publication.

It should be noted that within the media and think tank literature, blog posts represent an important source of references to PISA. As Farrell and Drezner (2008) show, blogs have an important first-mover advantage in influencing opinion. The ability to publish commentary quickly and cheaply makes blogs a key platform of information production

for both think tanks and the media. Figure 3.6 shows the proportions of blog posts and articles/reports in the sample. Especially noteworthy is that almost 60% of PISA references in *The Washington Post* come from its influential education blog, Answer Sheet, run by education columnist Valerie Strauss.

Figure 3.6: Sampled proportion of articles/reports to blog posts in think tanks and media



Data Analysis

To analyze the three sets of literature I conducted a content analysis of US education discourse. With this methodology, coding categories and explanations were produced directly from the emerging analysis of the data (Cresswell, 2007). The analysis was not a critical discourse analysis, which focuses specifically on linguistic practice, but rather a content analysis in the Lasswellian sense, looking for ‘who said what to whom, and why, and to what effect?’

Academic and think tank literature analysis

The analysis of the academic and think tank literature was a generative and iterative process. I first read all the articles in the sample at least twice, immersing myself in the data. These readings resulted in an initial list of ideas for broad topics that I could explore in the data. The initial list of possible categories included “arguments,” or how the article generally dealt with PISA and America’s performance; “externalization,” to see whether or not the article referenced a foreign example in its analysis of PISA (e.g., Finland); “policy borrowing,” referring to whether the article suggested the adoption of policy from other higher-ranked countries; and “actors,” to understand which actors are given voice in the discussion of US performance on PISA.

Initial categories were framed around the research questions, but these categories were revised over time in light of the ongoing analysis. This process of open coding allowed me to revisit and refine the categories as the research progressed. Analysis of and between the bodies of literature continued throughout the analytic process, as insights from one literature informed the analysis of the others.

Media analysis

The media analysis used a slightly different protocol than the academic and think tank literature. Findings from the initial qualitative analysis of academic and think tank literature informed the initial reading of the media articles. As with the academic and

think tank literature, I first immersed myself in the media data, reading the entire sample of articles at least twice. I began to develop coding categories aligned with the academic and think tank analysis, where applicable. Given the breadth of articles from the media and drawing on Pizmony-Levy (2013) and Olafsdottir (2007), I developed a plan for two-levels of analysis, at the article and the speaker level, which I describe in greater detail below. I used the concepts of framing and voice to help focus and deepen the analysis.

Article-level analysis. Articles were first coded for basic background information: by media source, author, date and headline. They were then coded by type, that is, whether they were news articles, editorials, or features from the print edition, or whether they were news articles or blog posts from an online edition of the source. Articles were also coded by location, noting the section of the paper in which they appeared (e.g., front page, national, international, business, editorial, features, or other). I also coded for any visuals (charts – with or without rankings, graphs, photographs, or other graphics) accompanying the article. Beyond the basic background information for each article, all coding was simple binary coding, meaning that each variable was coded as “1” when it was present in the article or “0” if it was absent.

To measure the objective of the article, I coded for the ‘news moment’ of the article: articles that were published within a one-month period surrounding the release of new PISA results; articles that discussed PISA scores but were published more than a month after the release of scores; and articles that referred to PISA, but were actually written to

discuss another topic. While this variable was designed to measure the subject matter, it also incorporated the notion of time, which would be important to the analysis later.

To measure the prominence of large-scale international assessments in the discourse, I coded as to whether articles mentioned PISA in a simple way or with a more detailed discussion, which I define as at least 20% of the article devoted to a discussion of the PISA assessment. I also coded for mentions of other large-scale international assessments, like TIMSS, PIRLS or CIVED, as well as mentions of the Common Core State Standards (CCSS). Though the CCSS are not an assessment per se, PISA was used as a benchmark against which the standards were developed (Carmichael, et. al, 2009). Additionally the Common Core, and more specifically the new standardized assessments that are attached to the Core, have become increasingly contentious in American education politics. Given that education elites used PISA as evidence to develop the Core, I use the CCSS as a testing ground for the externalization framework. Theory suggests that actors might reference PISA to justify contentious policy reform. Another question the research touches on is to what extent actors make increasing reference to PISA vis-à-vis the CCSS, as the issue becomes more contentious?

To measure the extent to which reference societies are used in the US discourse, I coded for references to other countries participating in PISA and whether the mention was simple or detailed, again defined as a discussion of 20% or more of the article. I also coded explanations for top-scoring countries' success. These explanations were broken

down into sixteen variables across four categories: economic, educational, cultural and political explanations.

I created three more main categories of variables to identify the “frames” of the discourse. Framing refers to the central organizing ideas of the discourse. How the media frames an issue determines which facts are deemed relevant, how thoughts are organized, and which arguments become key to the discussion. How a public policy issue gets framed in the media is particularly important in the US, where journalists play a stronger role in interpreting policy than in other countries, where political parties have more influence on the discourse (Halling and Mancini, in Ferree, et. al., 2002, 81).

I developed thirty-nine variables across the three categories to measure the strength of different frames. The first category was US education status, to measure how the articles connected PISA or PISA results to the status of education in the US. Within the category of general US education status, articles were coded more specifically for ideas such as the degree to which the US education system was presented as being in crisis, and what specific measures the article called for to improve US education. Additionally, there were a series of variables for articles that were skeptical of PISA and claimed that PISA had no bearing on US education status.

The second category of variables addressed claims connecting PISA results and US economic status. These encompassed any linking of PISA to US education performance

and economic competitiveness. There was also a category for those articles that claimed there were no links between US education status, PISA performance and economic competitiveness.

The third category measured the degree to which American performance on PISA was being linked to calls for policy action and/or educational reform. That is, were PISA results used to call for certain changes in policy? The variables included such issues as teacher pay, training, and evaluation; curriculum standards; school culture and other school-based reform measures. I also included a variable for calls for general reform as well as those articles that argued that PISA scores should have little or no effect on US education policy.

Speaker-level analysis. In addition to seeing how educational policy issues were framed in relation to PISA, I explore the notion of standing; that is, which actors were given a voice in the public debate. As with framing, standing is an essential component of policy discourse, as it determines which actors are given legitimacy in the debate. As Ferree, et al. (2002) note:

“[Standing] refers to gaining the status of a regular media source whose interpretations are directly quoted. Standing is not identical to receiving any sort of coverage or mention in the news; a group may appear when it is described or criticized but still have no opportunity to provide its own interpretation and

meaning to the event in which it is involved. Standing refers to a group being treated as an actor with voice, not merely as an object being discussed by others.”

(p. 13)

In order to understand which actors have voice in the discussions about the importance of PISA for US education policy, I coded every speaker that was directly quoted in any news article. The speakers were categorized into eight main groups of affiliations: government; the OECD; NGOs; universities; school leadership; teachers/teachers organizations; parents/parent organizations; and students/student organizations. The focus here is not on the journalists writing the articles, but rather on whom is given standing to comment on PISA within the article.

I then followed the same protocol described previously for the article level coding to code every speaker quote in the articles. Quotes were coded thematically for how they referenced PISA, Finland, China or other countries. If quotes included explanations of success, these explanations were coded into the same economic, educational, cultural and political categories that were used at the article level. Lastly, quotes were coded as to which frames they used. Again, following the coding protocol used at the article level, quotes were coded across the same thirty-nine variables in three major framing categories: status of US education, economic status, and implications for policy.

In addition to the qualitative analysis, media data were also recoded and analyzed using basic descriptive statistics. All data were managed in Excel, and statistical analyses were run using STATA 12.0.

Limitations

Similar to other social science research, this study is constrained by several limitations. First, it has the limitations of all small-N studies, most notably the constraints they provide on generalizability. The strength of small-N comparison is the “thickness” of the descriptions they can provide in looking deeply at a place in time, and can be more sensitive to the complexity of relationships and arrangements that may be closest to how the political world works (Coppedge, 1999). Small-N studies are generally not strong methods for generalizing causal arguments. While my study aims to provide a detailed account of the use of PISA evidence in the US during a specific time period, I do not provide any causal arguments nor do I suggest that the US results are generalizable to other countries participating in PISA.

Second, there is the issue of selection bias. In this work, I have chosen a discrete number of sources, but those choices bring inherent limits with them. For example, while the AERA journals may reasonably be considered a slice of the “American academic space,” there are many other journals published in the US that I might have chosen. Economics journals or other policy journals could provide a different view on the “American space.” This is true too in my choice of think tanks; other think tanks might have more extreme

partisan positions on education and would provide different data were they included. For the media analysis I have chosen high-quality, respected national American media sources, but my analysis does not explore coverage in local press or by wire services. In a sense, all of the data are drawn from a universe of elite sources, which introduces a particular orientation in the data. Each of these choices, while wholly defensible, points to some amount of bias inherent in the dataset.

One might also argue that my findings are historically shortsighted. Though the first decade of testing is an important marker in the history of PISA, it may turn out to be too recent to understand the ways in which PISA may ultimately affect the discussion of American education and education policy.²⁶ It is fair to assume that the engagement with ILSAs will continue to grow as the number of country participants continues to grow, as there is more data available, and, in the US, as the public becomes more familiar with PISA and as new political regimes come into power. Perhaps it is best to consider the findings as the initial framing and effects, which will evolve over time.

An additional point is specific to my coding. There may be some instances where the coding of two themes within an article does not express the emphasis given to one theme over another. That is, the coding may not always express the relative weight within articles of multiple themes. Also, though I acknowledge it here, my analysis does not

²⁶ Academics have engaged in debate regarding PISA's methods, validity and reliability of the outside the date range of the sample. In 2013-2014, for example, there were at least three public, politicized debates about the use of PISA data. See Carnoy and Rothstein (2013a; 2013b); Loveless (2014a); and OECD (2014b).

directly address the constraints placed on writers by space and by their editors, be they in the academic, think tank or media realms. In the media in particular, the desire for short, crisp news items may eclipse the ability for journalists to tell a more complex nuanced story; rankings may make a better headline than the philosophical and technical considerations raised by PISA.

CHAPTER FOUR:
What we talk about when we talk about PISA:
Findings from American educational discourse

In this chapter, I present findings from the analysis of PISA in American educational discourse from 2000 to 2012. The findings are drawn from an original sample of 130 texts from three separate literatures that can be conceived of as constructing American policy knowledge and creating a discursive American education policy space: academic journal articles, think tank publications and the media. Following a brief overview, I present findings from each literature in turn.

Across all literatures, I find that PISA has become incorporated into American educational discourse as a taken-for-granted source of evidence. In the academy, in think tanks and in the media, PISA is an object of curiosity and analysis, especially after 2009 when China first participates in the assessment and is ranked first in all subject areas.²⁷ Though there is some critique of PISA, primarily in the think tank literature, each literature embraces PISA as a model of assessment and as a measure of educational excellence. That is, there is a common understanding throughout the American education space that a country's performance on PISA is a representation of the strength of its education system. These understandings exist regardless of literature type or, in the case of think tanks and media, ideological leaning.

²⁷ Prior to 2009, Hong Kong and Macau had participated on PISA. The league tables list them as participating units, as Hong Kong-China and Macao-China. In the rest of this paper, as I discuss China's participation in PISA, I am referring to the participation of Shanghai, which participated as its own unit in PISA 2009, and was ranked first in the league tables in all subjects for that PISA cycle. When I discuss references to China, I am referring to references to Shanghai's participation. It is specifically Shanghai's participation that precipitates the change in American discourse that I detail in this chapter.

However, a close analysis also uncovers a series of inconsistencies in how PISA is used and understood. Most notably, even though PISA has been rapidly incorporated into American educational discourse, it has not resulted in concrete calls for lesson drawing or policy emulation. When in-depth discussions of top scorers do occur, they follow a “yes, but...” construction, where the success of top scorers is acknowledged, but then dismissed as coming at too high a cost (Tobin, in Cummings, 1989, 99). This is perplexing, given that the PISA discourse unfolds at precisely the moment that theory suggests that comparisons to elsewhere are most powerful, during highly contested domestic policy debates (e.g., the implementation of the Common Core State Standards). My analysis addresses this puzzle by tying comparative education theory more closely to the policy process, specifically to the ways in which problems are defined in domestic political discourse.

As I argue later on, rather than offering a new paradigm for thinking about domestic education issues, the PISA discourse employs educational ideas and discursive frames that have been dominant since the release of *A Nation at Risk*. Though the discourse is resonant, it has not been disruptive. PISA works to problematize the American education system by de-contextualizing the US and comparing it to other countries. However in order to name factors that might account for successful country performance, PISA needs to be re-contextualized in a national context. When this happens, explanations of PISA success turn out to be contextually specific, and not easily transferable. The data show that despite the forceful language of performance comparison, the pattern of

decontextualization and recontextualization has not served to promote change in policy. PISA data has not been used in the US to disrupt local agenda setting.

In the sections that follow, I examine in detail the ways in which each of the three literatures addresses PISA. I find that each of the literatures engages with PISA slightly differently, though commonalities exist. Academics use PISA primarily as a source of data for new lines of research, while the think tanks and the media see PISA as a measure of competitiveness, a reflection of a country's educational excellence. The academic literature remains largely self-contained, while there is frequent cross-pollination of ideas and experts between the think tanks and the media. Though the media drive the overall discourse, think tanks punch above their weight and are pivotal in introducing crisis language to the discussions.

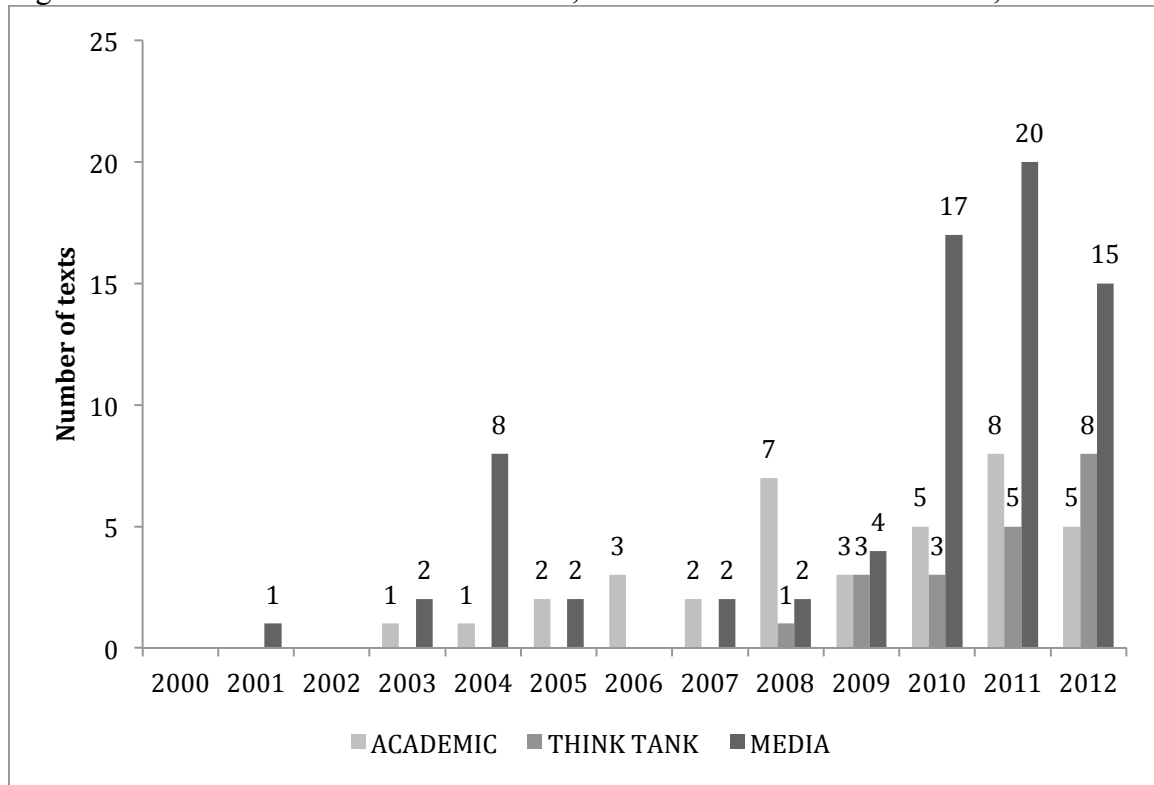
I also show how think tanks and the media use references to "top-scorers" as an important rhetorical device to problematize American education, and to energize and direct the discourse. The use of reference societies, specifically references to China and Finland, are a crucial part of the narrative, creating tension and distance between the US and other top performing countries. My analysis reveals that while PISA has gained a strong rhetorical foothold in American educational discourse, it is a profoundly elite enterprise, with extremely limited public voice. That is, those closest to the schoolhouse are given the least say.

Overview: PISA Across Literatures, 2000-2012

During its first decade, the status of PISA in the United States changed dramatically, from an unknown testing instrument to a bellwether of educational excellence. PISA achieved success on two fronts. First, it became widely accepted as a novel and model assessment instrument, one in which the emphasis is on applied knowledge rather than curricular recall. Second, it became a standard against which national education systems are measured. A country that ranks highly on the PISA league tables is considered to have a successful school system.

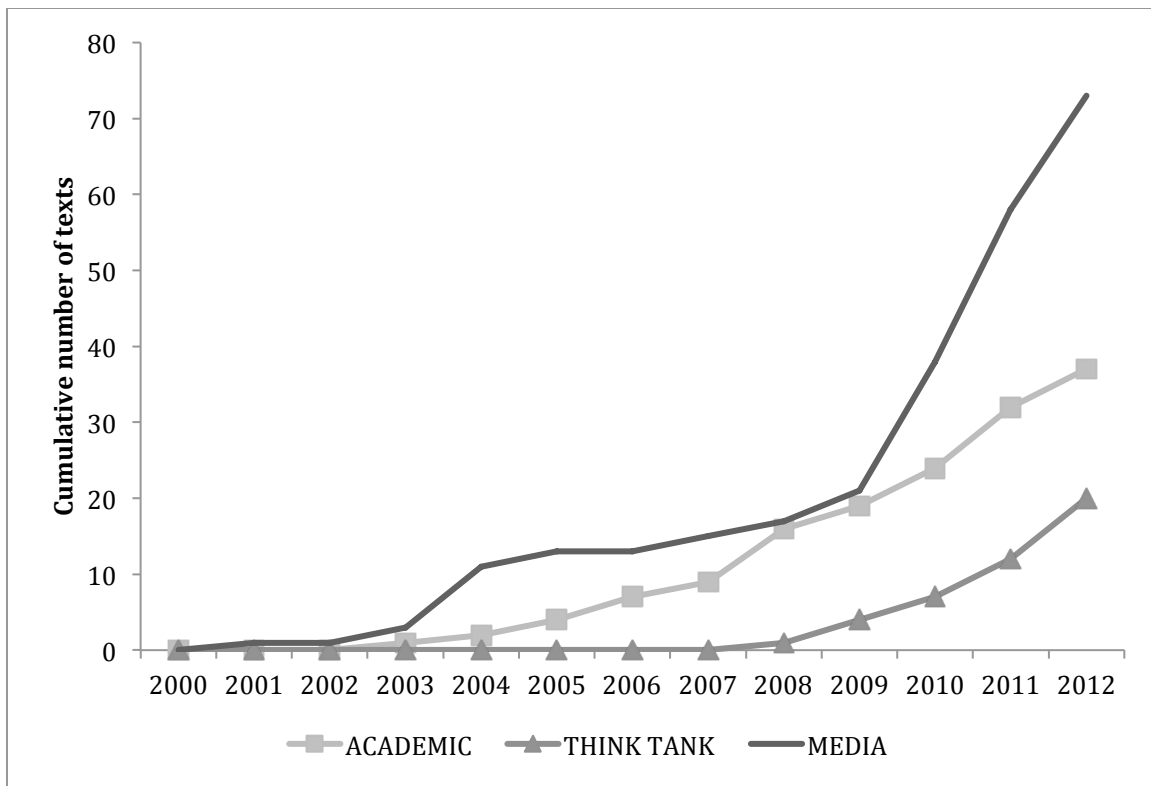
The dramatic incorporation of PISA into American educational discourse is shown in Figure 4.1, which charts references to PISA across the three literatures from 2000-2012. As I discuss in the rest of this chapter, each of the literatures engages with PISA data differently, yet each embraces PISA as an important source of evidence in education policy debate.

Figure 4.1: References to PISA in academic, think tank and media literatures, 2000-2012



I frame the discourse by first focusing on its evolution over time. Figure 4.2 shows the discourse unfolding in two phases via the cumulative publication of PISA literature over time. During the Phase I, from 2000-2009, PISA becomes a topic of interest. There is a slow but steady increase in references to PISA across the academic literature and the media, with the think tank literature joining the discourse slightly later, first referencing PISA in 2008. In 2004 there is a spike in references to PISA, primarily in the press, owing to the coincidence of PISA and TIMSS scores being released simultaneously. The media provide more coverage of this round of testing than in the previous round in 2001.

Figure 4.2: Cumulative references to PISA, 2000-2012



Discussions of PISA level out until the discourse explodes in Phase II, between 2010-2012. I discuss the reasons behind this rapid growth later in the chapter. However it is worth noting that during Phase II, references to PISA shift not only in quantity but also in their placement in the news cycle. In Phase I, PISA is discussed as a news story, in direct reaction to the release of triennial results. Later on, PISA is covered not only when results are released but steadily throughout the years between release cycles, suggesting that PISA moves from being an intermittent news phenomenon to a taken-for-granted form of evidence in the larger American educational discourse.

To understand how PISA achieves this taken-for-granted status, the rest of the chapter focuses on each body of literature separately. Though variation between literatures is highlighted, I also analyze the texts across a set of common dimensions: the ways in which each literature uses PISA as a source of data and a source of evidence; the role that reference societies play in American discourse, focusing on the two most frequently referenced countries (Finland and China); and whether PISA results are used as prescriptions for policy action.

Academic Literature

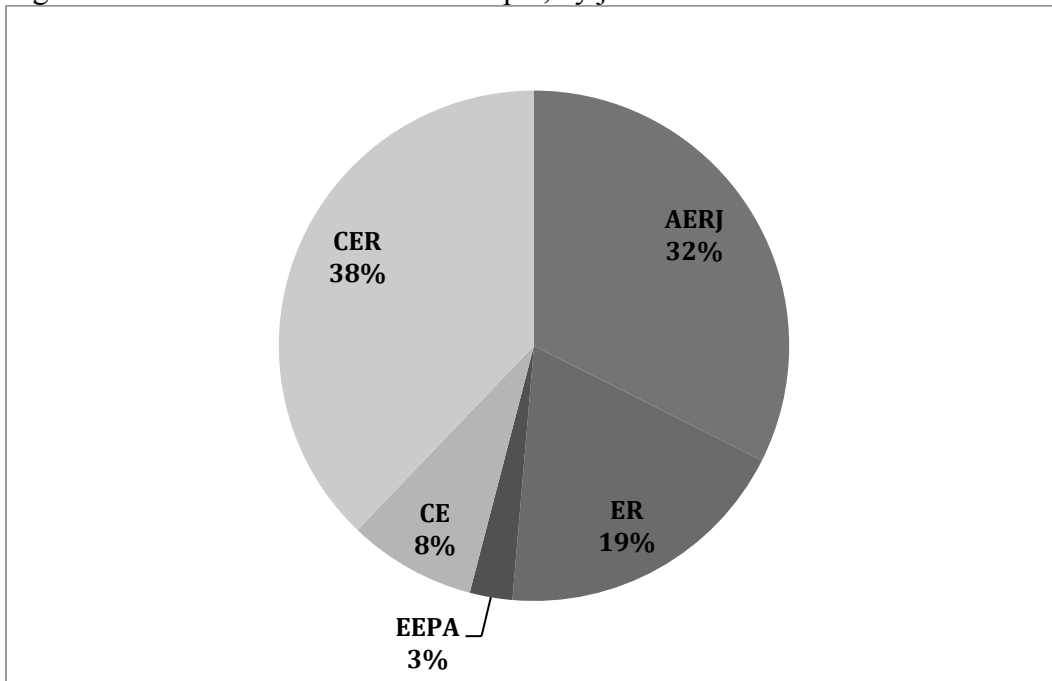
I begin by presenting findings from the academic sample, drawn from two sets of peer-reviewed, academic journals. In order to assess how PISA is understood in the American education space, I sample journal articles from three education journals published by the American Educational Research Association (AERA).²⁸ The search returned a total of 20 articles from the three AERA journals with more than half the articles coming from the American Educational Research Journal (AERJ), about a third of the articles from Educational Researcher (ER), and one article from Educational Evaluation and Policy Analysis (EEPA). Fifty-four percent of the academic sample is from the American literature.

Additionally, the academic sample includes 17 articles from the two leading journals in comparative education, Comparative Education Review (CER) and Comparative

²⁸ Though I use articles published by the AERA to constitute an American education space, not all of the authors themselves American, e.g. Kramarski and Mevarech (2005), who are Israeli.

Education (CE). Within the comparative education sample, the representation of articles is skewed toward Comparative Education Review (CER), which publishes articles that reference PISA about four times more often than the articles in Comparative Education (CE). Figure 4.3 provides an overview of the academic sample.

Figure 4.3: Overview of academic sample, by journal



Three main findings emerge from the American academic literature. First, the literature uses PISA primarily as a reliable data source for secondary analysis. This secondary analysis is on a variety of topics, though generally not topics related to the politics of PISA. The American academic literature develops a narrative in which PISA is used as a proxy for educational excellence, both in terms of ranking and as a model instrument of assessment. None of the American articles offers a critical look at the use of PISA.

These understandings of PISA are reflected in the comparative literature as well, though the comparative texts offer some critique of PISA.

Secondly, I find that the use of reference societies in the academic literature is minimal. Substantive²⁹ discussion of Finnish success is non-existent in the American academic literature, and Finnish success is only marginally discussed in the comparative literature. The same holds true for references to China, where only two references are made to China in the American literature and none are made in the comparative literature. Finally, though some academic literature suggests that foreign countries might be models for US education, the academic literature does not call for any specific lesson drawing or policy borrowing on the basis of PISA results.

PISA as a Data Source

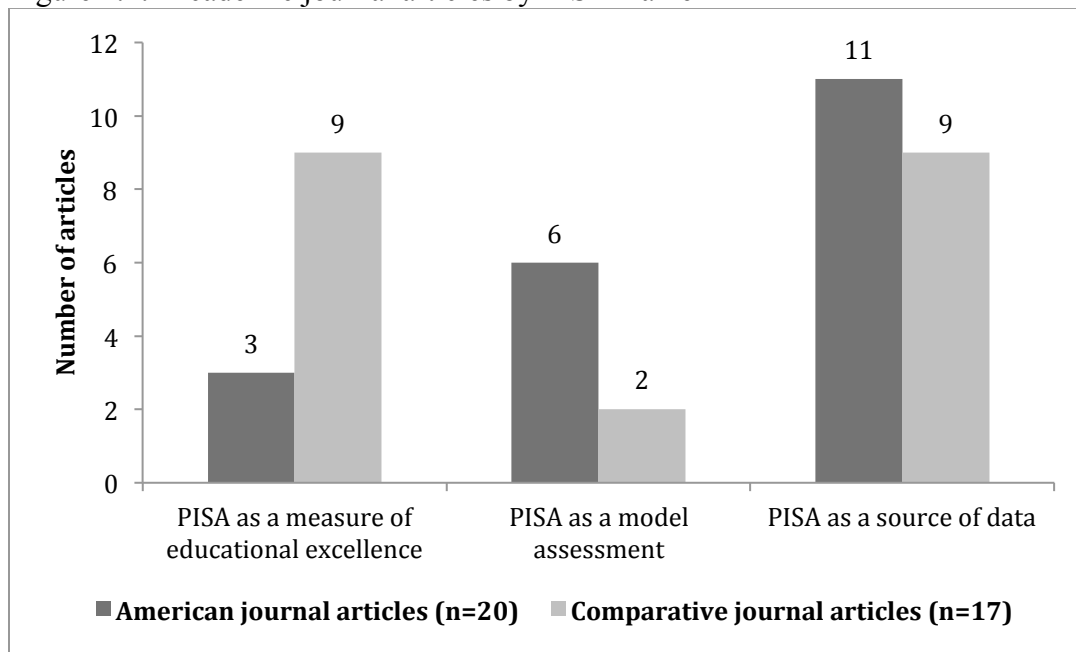
Looking at PISA representation in academic journal articles, I find that American and comparative academic texts reference PISA through three main themes (see Figure 4.4): (1) as a rich and reliable source of data (2) as a model assessment instrument and (3) as a measure of educational excellence. Both literatures consistently reference PISA through these themes.

Academics first and foremost use PISA as a valid, reliable and rich source of data. The growing number of countries, both OECD and non-OECD, that participated in PISA during its first decade has meant a growing dataset with significant variation. As

²⁹ In this paper, 'substantive' is defined as being the topic of interest in at least 20% of the article.

Rutkowski et al. (2010) point out, as more data become available through PISA from increased country participation and increased data collection the research opportunities also increase. Most academics agree that PISA test instruments have been rigorously and collaboratively constructed, and can be considered reliable, producing useable data from dozens of national education systems. Publicly available data analysis systems (such as the US Department of Education's International Data Explorer, or the OECD's PISA datasets) provide additional convenience. The academic literature describes PISA, as well as other prominent international assessments such as TIMSS and PIRLS, as unique sources of data that can support new lines of research.

Figure 4.4: Academic journal articles by PISA frame



One natural focus of research is on factors that affect student performance on PISA. Chiu and Khoo (2005) use PISA 2000 data to examine how resources, distribution inequality,

and privileged student bias affect academic performance. Luyten et al. (2008) use PISA 2000 data to assess the effect of one year's schooling for 15 year olds in England on reading performance. Condron (2011) uses PISA 2006 data to focus on the relationship between income inequality and achievement in twenty-seven OECD countries. In each of these studies scholars use multilevel regression analysis to examine how a series of independent variables might affect PISA scores.

Additionally, PISA solicits a trove of background information on individual, family and school level attributes. This allows scholars to explore not only factors that influence performance variation as measured by PISA scores, but also to explore topics that take variables other than performance scores as their outcome of interest. For instance, Seaton et al. (2010) use PISA 2003 data to investigate the big-fish-little-pond effect (BFLPE), the idea that students who attend high-achieving classes and schools have a lower self-concept than comparable students at middle- or low-achieving schools. In this instance, the authors use PISA data to address a psychological phenomenon that has little to do with assessment performance.

In the comparative literature, PISA is also used primarily as a source of data on a variety of topics, including the effects of various types of segregation on performance (Gorard and Smith, 2004); differences in gender on performance (Ma, 2008; McDaniel, 2010); the effects of regime type and sibship size on achievement (Xu, 2008); the effects of regime type and cultural capital on achievement (Xu and Hampden-Thompson, 2012); how

various tracking models effect achievement (Dupriez, Dumay and Vause, 2008); the effect of age-grade distribution on achievement (Dalton, 2012); and the relationships between parent-child communication, socioeconomic status, and levels of standardization in schooling (Park, 2008). The fact that such a breadth of topics is covered suggests that, like in the AERA sample, PISA is a growing source of rich, high-quality data for comparative academic work.

Limits to PISA Data

Several American texts point out limitations to the PISA data. These limitations address sampling and interpretation constraints, and are found in the standard limitations section of an academic journal article. Various sampling limitations are mentioned (Chudgar and Luschei, 2009; Chiu and Khoo, 2005; Rutkowski et al, 2010) and all authors note that the interpretation of data must be made cautiously, as the data is cross-sectional and not longitudinal. Language that implies causality cannot be used (Chiu and Khoo, 2005; Chudgar and Luschei, 2009; Seaton et al, 2010; Luyten et al, 2008). The comparative education literature also mentions these limitations (e.g., Park, 2008; McDaniel, 2010).

Rutkowski et.al. (2010) note that the technical background of the PISA data is complex and its two-stage sampling design diverges from that of other ILSAs. PISA follows other instruments in the first stage, by selecting schools on a probability-proportional-to-size basis, where larger schools have a higher probability of being chosen. However in the second stage, PISA randomly selects a set number of 15-year-old individual students

across various classrooms, whereas IEA studies randomly sample intact classes (Rutkowski, et. al., 2010, 143). The sampling design is not discussed in most articles, yet could be an important limitation to certain analyses. The articles that study the “big fish little pond effect” (BFLPE), for instance, use PISA data to test in part the idea that classmates are the most important variable in academic self-concept (Huguet et al., in Seaton et.al., 2009, 422). Yet whole classes are not sampled in PISA; students are sampled by age, across multiple classrooms. This suggests that PISA data may not actually be appropriate for these types of studies.

PISA as a Model of Assessment

The second way that PISA is referenced in the American academic literature is as a model of assessment. These references elevate the idea that PISA is an innovative instrument of assessment because it is testing the ‘real-world’ application of knowledge. They also implicitly support the idea that PISA scores are accurate cross-national measures of how well students will do in the ‘real world,’ i.e. in the job market. The references consider the non-curricular nature of PISA a strength, though none of the articles explains why a non-curricular approach is more appropriate or beneficial than a curriculum-based assessment. There is no critical examination of PISA as an assessment instrument. In fact, several scholars advocate for modeling US assessments on PISA and/or are reporting on interventions that are modeled on PISA. Yet some of these same scholars also call for a more critical approach to assessment.

For example, Stone et al. (2008) report on a professional development module they piloted and implemented to enhance the teaching of math to career and technical education students. The elements of their program “are very similar to the assessment framework that underlies the PISA, sponsored by the OECD” (Stone et al., 2008, 773). Another research team investigates the reading development skills of a cohort of students in British Columbia, beginning in kindergarten and ending in fourth grade. The students were administered an annual battery of diagnostic skills tests culminating in a standards-based comprehension assessment “comparable to that of other large scale assessments, such as PISA...” (Rupp and Lesaux, 2006, 323).

Baker (2007) calls for a more balanced approach to assessment in her 2007 Presidential Address to the AERA. She bemoans the shortfalls of an accountability movement that relies on testing, when:

“The evidential disconnect between test design and learning research is *no small thing* (italics in original). It means, at worst, that test may not actually be measuring the learning for which schools are responsible, thus gutting the basic tenet of the accountability compact.” (310)

Baker criticizes those assessments that are not designed to measure what schools are supposed to be teaching. Yet when she goes on to describe the assessment work that she and her team at UCLA are doing, she notes that they are using items from PISA as

models as they develop new middle school assessments. This is ironic, of course, because PISA is unaligned to any school curriculum and is precisely the kind of test she criticizes earlier. While one might think that academics would be the first to criticize PISA as a model assessment, they unexpectedly tend to embrace PISA as a model appropriate for assessment in the United States.

Leu et al., (2009) use this approach as well. In their research on understanding the teaching and learning potential of the Internet, they frame their findings in light of the importance of PISA:

“These findings are important as nations consider realigning public policies in education with the challenges of global competitiveness and information economies. State reading standards and state reading assessments in the United States, for example, have yet to include any online reading comprehension skills. This, despite the fact that several international assessments have already begun to do so, such as the Programme for International Student Assessment (PISA) and the Programme for the International Assessment of Adult Competencies (PIAAC).” (266)

Kramarski and Mevarech (2003) go even further. In their conclusion, they acknowledge the trend of requiring students to be able to explain their mathematical reasoning in

writing, and they note that teachers may want to alter their teaching to address the testing requirements:

“many state proficiency tests and international examinations (e.g. TIMSS-1999 and PISA, administered by OECD countries) include tasks that require students to explain their reasoning in writing. To acquaint students with such tasks and the scoring procedure, teachers may prepare guidelines and ask students to score one another’s explanations by using the guidelines...” (305)

The power of PISA as an internationally recognized model of assessment is evident. Even as academics caution against borrowing models from elsewhere, they embrace PISA with references that extol it as a model. Academics offer no critical view of the assessment itself or the politics of its use. However inadvertent it might be, academic silence about the limitations of PISA serves to legitimize PISA as a reliable and valid assessment instrument for US policy purposes. This silence also helps to develop a PISA tautology: the PISA assessment is good, therefore we should have more assessments like PISA.

PISA as a Measure of Educational Excellence

The third theme through which academics write about PISA is that of PISA as a measure of educational excellence. This theme is present in both the comparative and American academic literatures, and is slightly more frequent in the comparative sample. The

treatment of PISA as a litmus test of national educational excellence reflects the power that rank and comparison have as evidence among academic researchers. This framing is a straightforward one: the higher a country's PISA scores, the better its education system must be. In the American literature, references to rank on PISA are used to emphasize the performance of the US on the test, which is consistently characterized as mediocre to poor. None of the authors who reference PISA rank discuss any possible limitations of the ranking system, nor do they provide any caveats as to how the data could be misunderstood or misinterpreted.³⁰ Instead, the description of the US as substandard relative to other participating countries is how the authors' situate their research.

Schmidt and Houang (2012, 295) characterize US student performance on the math-focused PISA 2009 as "below par" as the US ranked 25th out of 34 OECD countries.

Vaughn et al. (2011) conducted a study to examine the effects of collaborative reading strategies for middle school students. They begin their study by noting that:

"There is not much room for optimism as students' reading achievement levels and understanding of complex text are at the lowest that they have been in a decade. The Programme for International Student Assessment (PISA) indicated that among the 30 countries that constitute the Organization for Economic Cooperation and Development (OCED)(stet), the United States ranked 21st on literacy-related outcomes. It is likely that a significant challenge for US students

³⁰ For a discussion of these issues, see Loveless (2010) and Labaree (2014).

is understanding the complex texts utilized in PISA assessments.” (Vaughn, et al., 2011, 940)

Authors also note the role that PISA rank has played in justifying education reform in the US. Baker (2007) shows “how tests exert power” by mentioning that lackluster performance on early international assessment helped to drive the more recent move toward increased assessment in schools. Schmidt and Houang (2012, 294) argue that it was the downward trend that US showed on international assessment performance during the past two decades that led to the development of its new national curriculum standards. While cautioning against a policy borrowing approach that takes methods out of context, PISA rank has helped to identify education systems as “world class” or not (Hopmann, in Luke, 2011, 370). Luke (2011) describes his own use of PISA rank when making recommendations for education reform to the Australian Labor government. In particular, he recounts how he highlighted the success of reforms in Finland and Ontario to government leaders, arguing that these countries were exemplary because they were “OECD darlings” and high scorers on PISA (Luke, 2011, 372-3).

Comparative education articles also discuss PISA rank as a reflection of national educational excellence. In several articles, PISA is mentioned only once, but as a litmus test of achievement (e.g., Hanson, 2006; Pizmony-Levy, 2011; Robertson, 2012). The fact that PISA has become deeply embedded in a culture of international comparison is also emphasized (Kamens and McNeely, 2010; Steiner-Khamsi, 2010), again showing

how PISA is understood as a measure of excellence. The academic literature considers countries with high PISA rank as models of excellence, and these countries tend to become referenced success stories. I turn to the use of reference societies in the next section.

Academic Use of Reference Societies

Using reference societies in educational discourse is a rhetorical strategy. Sometimes it signals an effort at the actual borrowing of policy, models or programs from another context. At other times, the reference is an allusion to something external to the national education system to provide a discursive source of authority or legitimacy. In either case, the use of reference societies can be read as symbolic of preferences and politics.

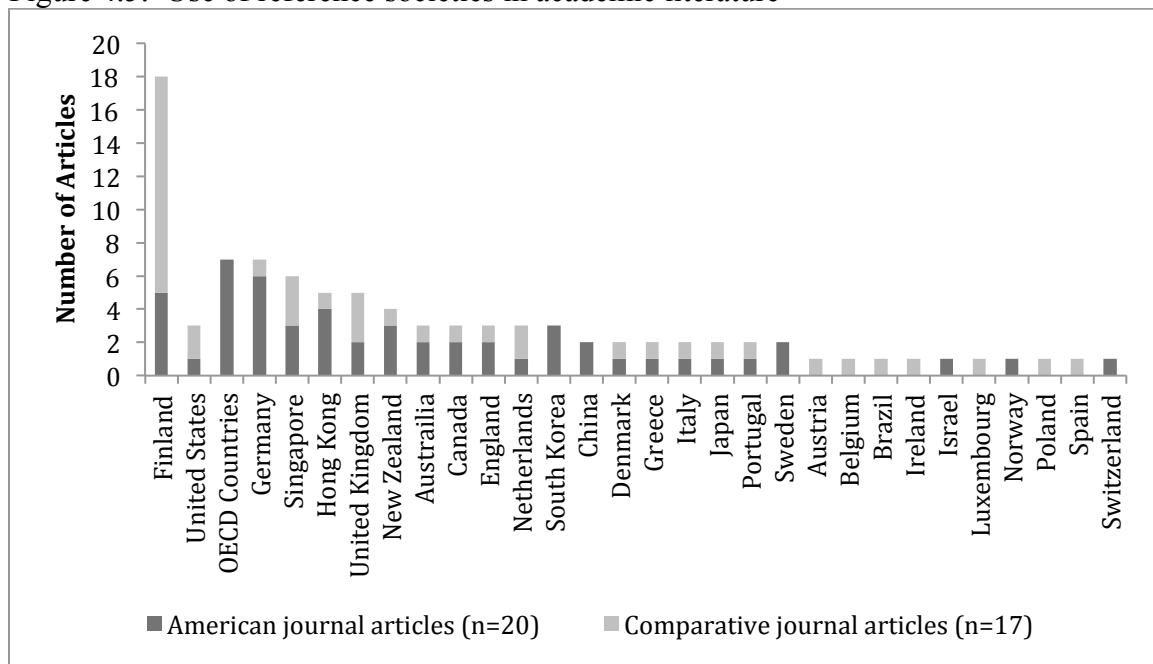
Analyzing which societies or standards are used as references provides insight into the ‘socio-logic’ (Schreier and Martinez, 2004) of the home country. Thus analyzing which countries’ PISA performances are referenced in the American discourse sheds light on the politics, priorities and agency of American education actors. I use this approach when analyzing reference societies in American PISA discourse.

In the American academic literature, a country is considered a reference society any time a country is mentioned as an educational model of some sort. Figure 4.5 provides details on which countries are used as reference societies in the academic literature. Finland emerges as the strongest reference society in the academic literature. As a consistently top-scoring country on PISA across all subject areas, educational elites have paid close

attention to the idea of “Finnish success.” A narrative of Finnish educational success developed from Finland’s consistently high rank and I follow the use of Finland as a reference society throughout the three literatures.

China is essentially unreferenced in the academic literature. This is noteworthy given the critical role references to China play in the think tank and media literature, as I show later on. China only participated in PISA for the first time in 2009, so part of the absence to references to China in the academic literature may have to do with the academic publishing cycle; results from PISA 2009 were released at the end of 2010, so that articles might not have had time to be written and published by the end of 2012. More than that, it suggests that American academics have ignored the political dimensions of PISA in favor of embracing PISA as a rich source of data.

Figure 4.5: Use of reference societies in academic literature



Finland in the American literature. Reference societies as a rhetorical device play a very small role in the academic discourse around PISA. While there are references to Finland in the academic literature, they gain little argumentative purchase. There are no references to China as a top scorer. Only a quarter of the articles reference Finland in the text at all, and none of those references is to Finland in any detail.³¹

The pattern in these articles is that Finland is briefly described as a “top scorer” and then referenced as a “model” in some way (Darling-Hammond, 2006; Condron, 2011; Luke, 2011; Chiu and Khoo, 2005; Baker, 2007). This is a hallmark of the entire PISA narrative. It is notable that all of the references to Finland, however brief or insubstantial, are positive. None of the articles questions any aspect of the Finnish system. While there is the occasional caveat that Finland is different than the United States in its history, demography and political system (e.g. Luke, 2011; Condron, 2011), the academic literature articulates Finland as a model of educational success.

Though each article discusses a different research question,³² all of the articles cite similar explanations for Finnish success. Finland is a model of equity and/or a model of

³¹ Four of the articles were published in *Educational Researcher* (Baker, 2007; Condron, 2011; Darling-Hammond, 2006; and Luke, 2011) and one in *American Journal of Educational Research* (Chiu and Khoo, 2005).

³² The five articles explore the following topics: Does privileged student bias affect a school’s or country’s overall academic performance? (Chiu and Khoo, 2005); In affluent countries, does economic inequality impact societies’ ability to produce very high or very low achievers? (Condron, 2011); How might the US create better balance between promoting common performance and supporting individual talents in assessment? (Baker, 2007); How well does policy cross borders, and what are the consequences of attempts to transfer policy? (Luke, 2006); What is needed to build a 21st century education system? (Darling-Hammond, 2006).

teacher preparation. For example, Condrón (2011) explores the connection between economic inequality and educational achievement. He notes that “top scoring Finland” has much less economic inequality than the US and suggests that this accounts for its high average achievement. He models US economic relationships under Finnish conditions, using Finnish levels of GDP per capita and income inequality to predict average achievement. By contrasting Finland and the US Condrón suggests that the US would do well to heed the importance of inequality on PISA performance. Though the author acknowledges that the countries “differ in many other ways that could shape achievement” (Condrón, 2011, 50- 51), he nonetheless decontextualizes Finland and the US to compare the two. Chiu and Khoo also reference Finland in this way, as one of the “top scoring countries in mathematics, reading and science...that pursue education policies that give each student equal funding” (Chiu and Khoo, 2005, 597).

Calling Finland an example of a “high quality, high-equity system,” Luke (2011) uses Finland to make the case that policy can ameliorate the impact of socioeconomic background on student achievement. There is some amount of irony in his argument which, while highlighting the limits of policy borrowing, nonetheless recalls his own failed attempts to get the Australian government to take a close look at Finland as it was undergoing education reform in the late 2000s (Luke, 2011, 372-373). While stating that the success of any model is culturally contingent, Luke references Finland, Ontario and Singapore – all high-scorers on PISA – as models of reform. Though he is sensitive to

the limits of policy borrowing, he nonetheless relies on comparison as a guide for his policy recommendations.

Finland's teacher training system is the second reference made in the literature.³³ Both Luke (2011) and Darling-Hammond (2006) use Finland as a model of an excellent teacher education system. Darling-Hammond in particular devotes a few paragraphs to discussing Finland's celebrated teacher education system, mentioning its dramatic overhaul in the early 1990s, its competitive selection process, reasonable pay, and strong content and pedagogical preparation. The text ends by linking this strategy of teacher preparation to equity, arguing that

“Finland.... sees no advantage in constructing a fundamentally unequal system in which a large share of the teaching force is poorly prepared and assigned to educate the most needy students, whose schools are routinely underfunded”
(Darling-Hammond, 2006, 20).

Thus while Finland is referenced only minimally, and never in the context of an in-depth discussion of Finnish success, the references are essentially the same across the texts.

Finland represents a strong investment in teachers and in equity.

³³ Finland's teachers are referenced across literatures as the reason for Finnish success, as I will demonstrate as this paper progresses. Ironically, Finland's teacher education policies are completely different than those promoted by the OECD. As Robertson points out, “In the OECD's PISA rankings, Finland ranks number one for student performance. The key elements the OECD proposes for teacher policies and high performing schools are absent in Finland...Finnish teachers spend fewer hours in class teaching than the OECD average, have considerable personal autonomy, are not engaged in formal systems of teacher evaluation and do not get merit pay” (Robertson, 2012, 596).

Finland in the comparative literature. Similar to the AERA sample, references to Finland in the comparative sample show little variation. References to Finland are considerably more frequent than in the American sample (thirteen out of seventeen articles), but as with the American texts, the substance of the references to Finland is minimal. Most simply refer to Finland as a top scorer, without further explanation or discussion. Finland is briefly referenced as a model for reform (Kamens and McNeely, 2010; Steiner-Khamsi, 2010), and once as a model of equity (Gorard and Smith, 2004) but generally Finland is referenced as a league leader. Whereas Finnish teachers are a primary focus in the American literature, the comparative literature only discusses teachers in one article, where it is ironically noted that the policies that the OECD proposes for teachers are absent in Finland (Robertson, 2012, 596).

The comparative education sample includes two articles that offer a more critical view of the idea of Finnish success. In these articles, the narrative of Finnish success is deconstructed into historical, cultural and political explanations that are context specific. Finnish scholar Hannu Simola analyzes the so-called “Finnish miracle” of success on PISA, and argues that Finland’s success owes to a “curious contingency of traditional and post-traditional tendencies” (Simola, 2005, 466). Simola’s point is that the Finnish system is particular, and to assume otherwise is to ignore the socio-historical legacies that gave rise to current education circumstances. Writing about Japanese education reform, Takayama (2010) unpacks the use of Finnish success in the Japanese context. He shows that both progressives and conservatives used reference societies and PISA rank to create

a political crisis in order to push for their preferred educational reforms. In Japan, the political left lionized Finland as a model for reform.

These critiques of Finnish success raise two important points. First, they show how Finland has become objectified and sensationalized through its PISA rank. As Grek (2009) has described, this is part of the de-territorialized, decontextualized ‘PISA effect,’ whereby context becomes irrelevant.³⁴ The fact that Finland is “top scoring” in and of itself legitimizes Finland as a point of reference. Most analyses of the Finnish “miracle” remain devoid of history or culture. In Japan, policy advocates created a vision of Finnish education that was purported to be exactly opposite of what was happening in Japan, thereby creating a crisis narrative that could drive support for reform (Takayama, 2010).

Secondly, the comparative articles show the political nature of the PISA discourse. The comparative education sample shows how Finnish success symbolizes an ideal of excellence as long as it remains unchallenged; Simola (2005) and Takayama (2010) both argue that when the concept of Finnish success is unpacked, its referential power is weakened because Finnish success then refers to a specific time, context and history. Left unexamined, Finnish success can mean all things to all people – it becomes “a canvas on to which the committed can project what they want to see” (Smithers, 2004,

³⁴ The OECD considers this a fundamental strength of PISA: “PISA offers a new approach to considering school outcomes, using as its evidence base the experiences of students across the world, rather than in the specific cultural context of a single country. The international context allows policy makers to question assumptions about the quality of their own country’s educational outcomes.” (OECD, 2001, 27)

iii). Especially in Japan, Finnish success is activated and gains traction at a moment in Japanese politics that reform ideas are contentious. The idealized Finland becomes a symbolic and discursive mechanism to push along the reform discourse (Takayama, 2010; 2013).

No Lesson Drawing or Calls for Emulation

The externalization framework suggests that references to foreign examples are used in order to legitimate contentious policy programs domestically (Waldow, 2012; Steiner-Khamisi, 2014). Despite the fact that the first decade of PISA coincided with contentious policy issues in the United States, references to PISA in the academic literature are not used to direct policy change. This is puzzling, since indicator data lends itself to simple comparisons about which systems are ‘better’ and ‘what works.’ There are no concrete calls for emulating Finland, China or other countries, even when those countries are referenced as some sort of educational model.

As the only author who addresses policy borrowing head-on, Luke (2006) discusses the possibilities for policy borrowing briefly, but by offering a round up of others who have suggested educational borrowing in their writing:

"Writing in the NYT, columnist Nicolas Kristof (2011) recently argued that America should look to China for examples of education reform. He praised the discipline and focus of Chinese teachers and students. Also in the NYT, globalization writer Thomas Friedman (2011) proclaimed the value of Singaporean mathematics education, although he cautioned of the need to borrow Singapore's pedagogy without its approach to individual freedom. In her important book *The Flat World and Education* (2008), Linda Darling-Hammond discusses Singapore and Finland as models for reform. Recently, US Secretary of

Education Arne Duncan convened a summit of OECD countries featuring discussions by representatives of school systems union -- with Ben Levin (2008) former deputy minister of education for Ontario, outlining Canadian reforms.... Could it be that American education is on the cusp of “outside-in” reform – the historical flows of expertise, innovation educational science and policy from the United States have reversed?" (Luke, 2011, 368)

Luke goes on to describe his own failed attempts at recommending Finland as a model for Australian policy makers to consider in their education reform efforts. Luke notes that the intervention failed because instead of borrowing reforms from ‘successful’ Finland, it borrowed the model that Joel Klein used in New York City. He then admits that any calls for borrowing are selective, and policy cannot be taken out of context (Luke, 2011, 375). He concludes that policy borrowing is historically and contextually bound, and must be grounded locally. Here again is a tension in the academic literature, simultaneously critiquing and accepting the notion of borrowing educational ideas from elsewhere. This is representative of the academic literature, ironically accepting of PISA and league leaders as models while critical of policy borrowing more generally.

Conclusion

This section presented findings from the academic literature. A total of thirty-seven peer-reviewed journal articles were sampled from three AERA journals and two leading comparative education journals. Across the academic literature, PISA is used primarily as a source of data for secondary analysis on a wide range of research questions. None of the academic literature I reviewed seriously challenges the presentation or legitimacy of the PISA data.

Despite PISA's stated purpose of being de-coupled from any specific national learning goals, several of the articles suggest PISA as a model for national-level assessments. The data show that in the US and other national contexts, PISA achievement scores are being used as evidence to drive assessment format and/or specific classroom level interventions. PISA rank is considered a measure of national educational excellence. Only two articles question the ways in which PISA is used as evidence in educational discourse.

Both the academic and comparative literature favor Finland as a reference society. China plays no role as a reference society in the academic texts. References to Finland are universally positive and use Finland as an example of high academic achievement. Finland is also seen as a model of equity and as having superlative teacher education system. Finland is referenced as a success story even in those articles that focus on the difficulty of transferring educational policy from one country to another. There is also no clear celebration of additional specific countries that have done well on PISA (e.g., Korea, Singapore).

Finally, the academic literature has no serious calls for policy borrowing based on PISA, even though PISA is understood as valuable because its rankings provide a list of who's 'best.' The academy regards PISA first and foremost as a data source, and uses it to pursue social scientific research. The American academic literature does not engage in critical discussions of the construction or use of PISA and steers clear of suggesting

borrowing any policy directions from other countries. While the academic research may be of high quality, its findings and interpretations are not brought to bear in the think tank or media literature.

Academics provide notably little critique of PISA. This silence serves to legitimate PISA as a model assessment tool and a measure of educational excellence. As a source of knowledge production, the academy “holds the power of interpretation” (Carney, Rapplee, and Silova, 2012, 384) and plays a key role in developing the normalizing discourse of international assessment. The fact that the American academic literature uncritically embraces PISA – as a source of data, as a model assessment, and as a measure of excellence – creates a space for PISA to become increasingly accepted as an important guidepost for educational policy.

Given the role of academics as analysts in defining educational problems and solutions, their mostly uncritical embrace of PISA is surprising. Even as academics caution against borrowing from elsewhere, they embrace PISA scores as real measures of comparative academic excellence. Universities are acting as “receptor sites” (Frank, et al., 2000) wherein academics play a central role in interpreting and disseminating educational knowledge. However in the case of PISA, the expert academic knowledge that is being generated has not offered any significant reinterpretation of PISA in a critical way. Instead, American academics are producing knowledge that reinforces PISA as a valid and usable source of data and a model of assessment.

Think Tank Literature

In this section, I present the findings from the think tank literature. These findings are drawn from data gathered from two prominent American think tanks, each of which has a strong focus on education research: the Thomas B. Fordham Institute (“Fordham”), a politically conservative education think tank based in Ohio and in Washington, DC, and the Brookings Institution (“Brookings”), a centrist/liberal think tank also based in Washington, DC. The think tank sample consists of twenty texts, closely divided between Fordham and Brookings (nine texts from Fordham and eleven from Brookings). Think tank texts are drawn from reports, papers, and blog posts from each think tank.

Think tanks have played an increasingly important role in political discourse during the past several decades, providing ideas, experts and data and analysis for political debate (Rich, 2004). As organizations, think tanks want to produce high quality research for use by policy makers, other analysts and journalists (McGann, 2013). Given these incentives, it is not surprising that education analysts at both Fordham and Brookings have put PISA on their institutional research and advocacy agendas; rankings are in many ways perfectly suited for simple, attention-grabbing accounts of winners and losers.

McDonald (2014) describes the move of conservative think tanks into the education policy field, which had been historically dominated by Democrats with a focus on equity politics. After *A Nation at Risk* was published, though, American education politics was radically restructured to a right-of-center paradigm, dominated by “both free-market

neoliberal policy prescriptions and a neoconservative push for government oversight in the form of assessment” (McDonald, 2014, 851). Reflecting this shift, recent accounts of American education politics discuss a collapse in traditional partisan approaches to education policy (Mehta, 2013; DeBray-Pelot and McGuinn, 2009; Tyack and Cuban, 1995). These accounts suggest that there has been more agreement than difference in liberal and conservative thinking on education in the decades since *A Nation at Risk* was published. One question, then, is to what degree political affiliation affects the way in which think tanks interpret American PISA performance. The answer, I find, is not much.

Fordham and Brookings have a similar narrative, with both think tanks celebrating PISA as a model of international assessment, even as they simultaneously provide scathing critiques as to why PISA should *not* be used for benchmarking. Both think tanks use reference societies to construct the discourse, as evidence and rhetorical firepower. China is crucial to both think tank narratives, and both think tanks problematize American PISA performance in terms of China. The introduction of China into the rhetoric changes both the quality and the quantity of the PISA discourse in both the think tanks and the media. Finally, as in the academic literature, the think tank literature makes virtually no calls for policy borrowing.

Think Tank References to PISA Over Time

Though PISA results were available in 2001, neither of the think tanks mentions PISA until the late 2000s. There is no discussion of PISA by either think tank prior to 2008, when Fordham publishes an initial blog piece discussing the idea of Finnish success. Think tank interest in PISA subsequently explodes, and virtually all of the texts from both Fordham and Brookings are published between 2009-2012. Both think tanks have similar growth trajectories and the rapid growth in think tank discourse mirrors the growth patterns in the academic literature. The post-2009 growth in discourse reflects the interest of US education analysts in China's PISA performance, reflecting the larger particular geopolitical relationship between the US and China. China's performance provides a new threat, through which American education performance is viewed.

Traction Through "Crisis Talk"

Prior to 2009, with the exception of a short Fordham blog post debunking the notion of Finnish success, think tanks do not really focus on PISA country comparisons. In fact, they do not focus much on PISA at all. A dramatic shift occurs as the 2009 PISA scores are released in December 2010. It is the first time that China participates in PISA, and China receives top scores in literacy, mathematics, and science. China's top scores reverberate throughout the think tank literature and immediately a new narrative appears. China's scores are framed as a threat to the United States; the language of the think tanks shifts into crisis mode; and think tank interest in PISA explodes, including not only references to China, but to Finland and to PISA more generally.

The use of crisis talk originates in the Fordham literature. The language is colorful, direct, and sometimes extreme in its description of PISA and the state of US education. Chester Finn, Jr., president of Fordham, whose blog post sets off a deluge of reaction to the 2009 results describes it thusly:

“Fifty three years after Sputnik caused an earthquake in American education by giving up reason to believe that the Soviet Union had surpassed us, China has delivered another shock. On math, reading, and science tests given to 15-year-olds in sixty-five countries last year, Shanghai’s teenagers topped every other jurisdiction in all three subjects – by a sweeping margin.” (Finn, 2010, A21)

Finn goes on to warn that even though only students from Shanghai were tested, “Americans would be making a big mistake to suppose that this Shanghai result is some sort of aberration.” America, he notes, is in the middle of the pack in terms of PISA rankings and has been such for a long time. But:

“until this week we could at least pretend that China wasn’t one of those countries that was a threat....Today that comfort has been stripped away. We must face the fact that China is bent on surpassing us, and everyone else, in education and acknowledge what the consequences of this may be.

Will this news be the wake-up call that America needs to get serious about educational achievement?" (Finn, 2010, A21)

Here, China is threatening, competitive, and aggressive; to ignore its PISA success is to be at risk. The tone is alarmist as Finn makes the case for a fearsome China. The dramatic narrative is compelling, and it is redeployed in several other Fordham texts. Petrilli and Scull (2011) use the "Sputnik moment," and China's success is said to have "rocked the education policy community" (Fairchild, 2012). It is the thing that "everyone is talking about" (Joch, 2010).³⁵

The Rhetoric of Sputnik and China

The Sputnik metaphor is particularly powerful because of its historic connotations. The impact that the Sputnik satellite launch by the Soviet Union in 1957 had on the US – both materially and psychologically – was enormous. The launch caught the American public off-guard and politicians were able to capitalize on that vulnerability to accelerate the space and arms races.³⁶ To many, the Sputnik launch symbolized the failure of the US education system, that the United States wasn't doing a good enough job in educating its

³⁵ As I discuss in the section on media, President Obama and several other analysts and journalists subsequently use the Sputnik metaphor.

³⁶ The Sputnik launch was a turning point on many fronts. It led to the passage of the National Defense Education Act, which supported students to go into math and science education, the creation of NASA, and the creation of the Advanced Research Projects Agency (ARPA), the precursor to DARPA, the agency that developed the Internet and other advanced government technology. In education, the Sputnik crisis followed a "crisis" script that has characterized subsequent reform efforts: (1) a geopolitical rival of the United States outperforms the US in some way; (2) rhetorical hysteria ensues, and the education system is blamed; (3) education reformers, building on ideas in education reform that were already fomenting at the time, capitalize on the symbolism of the event and push a new era of reform. Bybee (1997) offers an historical contextualization of this pattern during the Sputnik crisis.

children especially in math and science. Subsequent analyses of the American system concluded that the US was far behind the Soviet Union in terms of subject content, sequencing, quality of learning materials and rigor of the curriculum.³⁷ The rhetoric of the Sputnik crisis linked education to national security, national pride, geopolitics and technical innovation, and began to strengthen the hand of the federal government in education policy and educational funding (e.g., Tyack and Cuban, 1995; Mehta, 2006).

Finn's Sputnik metaphor is picked up by several news outlets and crosses party lines. First, the blog post runs as an op-ed in the conservative Wall Street Journal. Then Democratic President Barak Obama picks up the Sputnik language. In a speech on the economy that the president gives in North Carolina within a few days of the release of results, Obama says, "our generation's Sputnik moment is back."³⁸ US Secretary of Education Arne Duncan echoes these remarks saying, "Unfortunately, the 2009 PISA results show that American students are poorly prepared to compete in today's

³⁷ See, for example, Trace (1962) and Bybee (1997).

³⁸ Obama discusses American economic competitiveness in terms of competition with China, India and South Korea, and focuses on the need for strong American education, saying, "We are the home of the world's best universities, the best research facilities, the most brilliant scientists, the brightest minds, some of the hardest-working, most entrepreneurial people on Earth -- right here in America... But as it stands right now, the hard truth is this: In the race for the future, America is in danger of falling behind. That's just the truth. And when -- if you hear a politician say it's not, they're not paying attention... In 1957, just before this college opened, the Soviet Union beat us into space by launching a satellite known as Sputnik. And that was a wake-up call that caused the United States to boost our investment in innovation and education -- particularly in math and science. And as a result, once we put our minds to it, once we got focused, once we got unified, not only did we surpass the Soviets, we developed new American technologies, industries, and jobs. So 50 years later, our generation's Sputnik moment is back. This is our moment." (Obama, 2010)

knowledge economy (Duncan, 2010).³⁹ New York Times' op-ed columnist Nicholas Kristof also uses Sputnik to describe China's top scores on PISA, noting

“Americans think of China's strategic challenge in terms of, say, the new Chinese stealth fighter aircraft. But the real challenge is the rise of China's education system ... We're not going to become Confucians, but we can elevate education on our list of priorities... That's what we did in 1957 after the Soviet Union launched Sputnik. These latest test results should be our 21st-century Sputnik.”
(Kristof, 2011, WK10)

Suddenly, PISA scores are symbolic of an education crisis. China's success problematizes the US system. Like the Soviet launch of Sputnik, China's top scores on PISA symbolize the US being bested by a geopolitical rival. The problem definition changes dramatically and US education is in crisis. Though the message is less about national security than economic competitiveness, unless the US raises its lackluster performance, it will be beaten.

The Sputnik metaphor marks a turning point in the discourse. In the think tanks and as I show later in the media, references to China energize and direct the discourse with a new sense of urgency. The crisis language in the think tanks around China's performance

³⁹ Indeed, Duncan asks the OECD to prepare a special report on what the US can learn from other high-performing countries, resulting in the publication of “Lessons from PISA for the United States: Strong Performers and Successful reformers in Education” Paris: OECD Publishing.
<http://dx.dio.org/10.1787/9789264096660-en>.

becomes the jumping off point for the post-2009 PISA discourse.⁴⁰ PISA becomes a measure of national educational excellence, even as both Fordham and Brookings offer significant critiques about using PISA performance as a measure of system success.

PISA as a Measure of American Educational Excellence

The PISA narrative in the think tank literature develops rapidly after 2009, when PISA is suddenly referenced as *the* measure of educational excellence and as a model assessment. The relative prominence of frames throughout the think tank literature is given in Figure 4.6. Both think tanks use PISA rank as a litmus test for educational success. Indeed, the titles of two of Fordham's major reports from this period reflect the importance given to international rankings (e.g., "American Achievement in International Perspective," (Petrilli and Scull, 2011) and "Stars by Which to Navigate? Scanning National and International Education Standards in 2009," (Carmichael, et al, 2009)). Fordham publishes three major reports between 2009-2011 that base their analyses on data from international rankings, concentrating on PISA. In the Brookings annual education reports in 2009 and in 2011, fully one third of each report is on PISA. By making PISA the focus of these publications and embracing PISA scores as a measure of educational

⁴⁰ Though not within the sampled articles, a contentious debate about Chinese participation has arisen between Andreas Schleicher and several PISA analysts. The Economic Policy Institute published a report in January 2013 titled, "What Do International Tests Really Show about US Student Performance?" by Martin Carnoy and Richard Rothstein that re-analyzed PISA 2009 data and took issue with characterizations of US performance. Schleicher responded, and Carnoy and Rothstein parried back. See Carnoy and Rothstein (2013a; 2013b). Brookings analyst Tom Loveless engaged in a pointed back-and-forth with Schleicher regarding the selective nature of Shanghai's school enrollment policies that force certain students out of Shanghai schools thereby skewing the sample. See Loveless (2014a) for an overview, and Footnote 1 of that same report for links to Loveless' previous commentary on the China problem.

excellence, both think tanks support the characterization of PISA as a litmus test for success.

PISA as a Measure of Foreign Educational Excellence

The think tanks also use PISA performance to measure success not only in the United States, but also in other countries. In this comparative approach, PISA becomes a normative measure of what educational success should look like in the developing world. Several texts about Latin America, for instance, make claims about the state of crisis in education as evidenced by poor PISA scores. One article notes:

“The 2009 OECD PISA evaluations on education quality around the world indicate that none of the Latin American countries surveyed are at or above the world average in reading, mathematics or science, and most are significantly lower than the average. Narrowing the gaps in enrollment across population, especially in secondary and tertiary education, and improving quality across the board should be first priority.” (Cardenas, 2010)

This type of reference is found in another Brookings report on the use of private funding to support public goods in Latin America:

“Several international and domestic assessments and indicators highlight the extreme nature of the learning crisis in the region. The most recent OECD

Programme for International Student Assessment (PISA) study measuring reading, science and math skills of fifteen year olds around the world ranked Chile the highest scoring country in the region at 44th out of 65 countries...” (van Fleet and Zinny, 2012, 1)

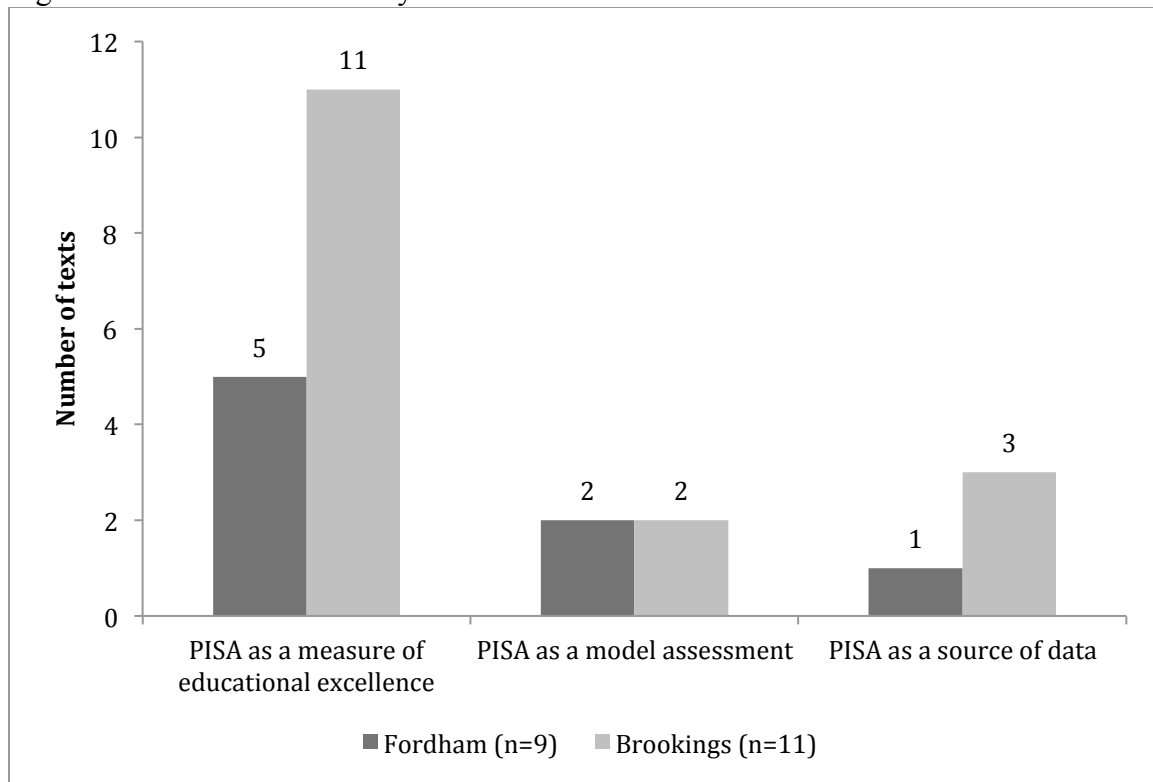
Brookings also publishes a report summarizing the proceedings of the Sixth Summit of the Americas in 2012, criticizing the participants for giving short shrift to education. As the authors write:

“...the reality is that most Latin American schools...provide low-quality education that does not meet countries’ or students’ needs. For example, roughly half or more of Latin American students participating in the OECD’s most recent Programme for International Student Assessment (PISA) exam performed at the lowest levels, meaning that they had difficulty applying basic reading and math skills to real world problems.” (Puryear and Goodspeed, 2012, 35)

American think tank analysts, writing about foreign countries with wildly varying education systems, make casual reference to those countries’ poor PISA performance as evidence of a weak state, from Mexico to the UAE. PISA is seen to reflect educational status in the Middle East (Kraetsch and Constant, 2010) and Qatar and the United Arab Emirates, where PISA is described as “the universally recognized measure of national education performance” (Bungalawala, 2011, 8). Ironically, often the think tank reports

are not even about education. These references are decontextualized, de-territorialized, simple and uncritical. Yet they show the degree to which analysts have incorporated PISA into their own repertoires of evidence, citing PISA rank in a wide variety of country settings.

Figure 4.6: Think tank texts by frame



Think Tank Critique of PISA

It is notable that while think tanks reference PISA as a litmus test of educational success, they simultaneously provide scathing critiques as to why PISA should *not* be used for benchmarking. Twenty percent of the articles that reference PISA do so by offering a critique. Liberals and conservative criticisms are the same across think tanks, with each

think tank publishing two pieces detailing PISA's weaknesses. The critiques are non-partisan; in fact, one of the Fordham texts praises the Brookings report for its "stinging indictment" of aspects of the PISA enterprise and encourages Fordham readers to read the Brookings text (Northrup, 2009).

The critics level several charges. The earliest critique is the 2008 annual education report from Brookings, released in 2009.⁴¹ The 2008 Brookings report states that while "PISA is influential in Europe... it flies below the radar in the United States" (Loveless, 2009, 8). In the context of increased calls in the United States to use PISA as a benchmark for the Common Core standards, the report cautions that "serious reform" is needed on PISA for several reasons.

The report charges that the OECD is making recommendations beyond what the data can reasonably conclude, using cross-sectional data to make causal claims. For instance, even as the OECD authors warn against confounding correlation with causality, they argue that PISA data show that those students who feel confident in their science ability perform better on science assessments. This is based on a positive within-state correlation between self-efficacy and science performance on the 2006 PISA. The

⁴¹ Tom Loveless, a Senior Fellow at Brookings' Brown Center on Education Policy, authored both the 2008 and 2010 Brookings education reports. Loveless was formerly the US Representative to the General Assembly at the International Association for the Evaluation of Educational Attainment (IEA). Amongst other assessments, the IEA administers TIMSS, sometimes seen as a kind of competitor to PISA, though PISA has arguably become more prominent in public discourse since 2009.

authors write “this strong association shows that building students’ confidence in their ability to tackle scientific problems is an important part of improving science performance” (OECD, in Loveless, 2009, 12). However, the nature of the data does not allow for such causal conclusions. The OECD’s PISA policy recommendations go beyond what the data actually support.

The Brookings report is highlighted on a Fordham blog post almost immediately after the report is published. Fordham encourages readers to take the Brookings critique seriously and then takes its own objections further, writing:

“Who cares either way? This isn’t a test of confidence. Its supposed to be about what kids actually know about math and science content, not what they think they know about math and science, or what they are motivated to want to learn!...Here’s hoping that we slow down and take a serious look at what this test is measuring before we deem it worthy of an international benchmark.”

(Northrup, 2009)

Both think tanks charge that PISA is ideologically biased. The 2008 Brookings report analyzes a battery of questions from PISA 2006, which focused on science. The analysis shows that on environmental issues, PISA questions assume certain normative definitions of responsibility and understanding, and score students on attitudinal beliefs instead of

their ability to apply scientific knowledge, evaluate hypotheses, or use evidence. The think tanks admonish PISA to just “stick to the science” (Loveless, 2009, 18).

Think Tank Rhetoric

Unlike the social-scientific, dispassionate tone of the academic literature, the think tank analyses are both opinionated and colorful. PISA performance is framed in terms of international competition, continuing the discursive logic begun in the early 1980s with *A Nation at Risk*. American performance is described as “lackluster” (Petrilli and Scull, 2012, 2), “lagging” and “depressing” (Winkler et al., 2009, 6), “disappointing” (Greenstone and Looney, 2011) and “inadequate” (Haskins et al., 2012, 2). The Olympics are used as a favored analogy; they become the model for one of Fordham’s 2009 publications, “Education Olympics: The Games in Review.” This report provides a detailed analysis of American educational performance compared with other top scoring countries on both PISA and TIMSS. Those countries that do best in various ‘events’ are awarded ‘medals’ by Fordham (Winkler, et al., 2008). Another report uses PISA data to examine American performance in comparative perspective, particularly how high and low achieving American students compare to similar groups of students in other countries (Petrilli and Scull, 2008).

Think Tank Politics: Linking PISA to the Common Core State Standards

The introduction and implementation of the Common Core State Standards throughout the 2000s would seem to have offered an excellent opportunity for think tanks (and other

education actors) to externalize to foreign examples. Many top scoring countries on PISA have national standards,⁴² and theory suggests that the contested nature of the CCSS would make use the foreign examples likely. Fordham in particular is a long-time supporter of standards and had begun to link PISA to the CCSS in some of its publications (e.g., Carmichael, et al., 2009; Schmidt, Hoang and Shakrani, 2009).

Unexpectedly, both think tanks come out strongly opposed to the idea that PISA should be used as a model for standards and assessment in the United States. Not only are both think tanks in agreement, but also they both dedicate significant publication space to their critiques (a major report, plus additional blog posts from Fordham, and a major report and additional article from Brookings). Both express concern that a consortium of state education organizations forms an advisory group to figure out how to benchmark state standards to PISA. The Co-chair of the NGA, Georgia Governor Sonny Perdue is quoted as saying:

“As governors, we must have consistent, comparable data in order to make informed decisions about our state’s education system [stet]. Benchmarking will help us identify the qualities and characteristics that make up the education systems that best prepare students for success. Understanding these policies give us the option of incorporating the best of them into our own educational structure.” (Loveless, 2009, 8)

⁴² Singapore, South Korea, and Canada, for example. For additional analysis see Schmidt, Houang and Shakrani (2009).

Fordham issues a major report on benchmarking as part of the public response period for the 2009 draft of the Common Core State Standards released by the NGA and CCSSO. Fordham reviews the draft Common Core standards along with NAEP, PISA and TIMSS in order to “help US educators and policymakers to judge the respective merits of these influential standards” (Carmichael, et al., 2009, 1).⁴³ The Fordham report assesses:

“the content, rigor and clarity of the first public drafts of the CCSS and – applying identical criteria – also to appraise these elements in America’s current de facto national math and reading/writing tests (NAEP) as well as two influential international testing regimes (PISA and TIMSS) that many look to for “benchmarking” purposes....Our goal is to help readers and users determine which of these document ought to influence their thinking, their standards, their aspirations, and their tests. Is one of these national/international frameworks worthier of emulation than others?” (Carmichael, et al., 2009, 7)

Fordham then goes on to slam the notion that PISA could be an assessment on which the CCSS should be based. In the Fordham report, subject matter experts reviewing the CCSS, NAEP, PISA and TIMSS, gave PISA a grade of “D” in both mathematics and reading, the lowest scores amongst the assessments reviewed, and opine that PISA items

⁴³ Fordham argues that while PISA is not a set of content standards per se, it de facto acts as a set of standards by offering prescribed content that it deems important for success in adult life (Carmichael, et al., 2009, 9).

should not be “a suitable model for US standard setters at any level (Carmichael, et al., 2009, 2).

Using common metrics to review each set of frameworks, the reviewers analyzed both the mathematics and reading literacy frameworks. The report finds PISA “unbalanced,” “weak,” “confusing,” and “murky” and concludes that

“PISA strikes out. Neither in reading (literacy) nor in math does its content deserve better than a grade of “D.” This is not a promising benchmark for American K-12 education in these subjects.” (Carmichael, et al., 2009, 6).

Brookings also issues a report that provides strong claims *against* PISA as a model for content standards or assessment (Loveless, 2012).⁴⁴

Different Ideologies, Similar Claims

Thus emerges a pattern where think tank ideology does not define political position. Instead, both think tanks simultaneously critique and embrace PISA. They both advocate against using PISA as a benchmarking tool for US educational standards, but nonetheless assess PISA as an important measure of educational excellence throughout their literatures. One explanation for this inconsistent use this may be institutional, at least

⁴⁴ Though it is not possible to draw a clear conclusion from the empirical evidence here, there are no additional efforts, notably in any of the media articles, to link the CCSS to PISA or other international benchmarks after these reports are released. Perhaps the stinging critique from Fordham and Brookings helped to push the NGA and CCSSO to change their rhetoric to de-link the CCSS from PISA or other international benchmarks.

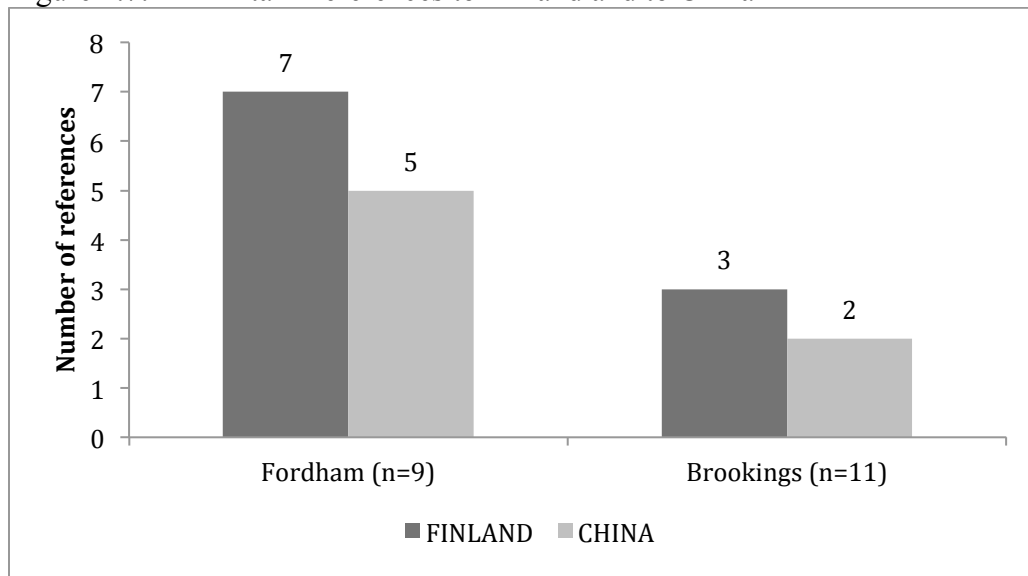
with regard to the Brookings analyses, which have this inconsistency. Given that think tank analysts may operate with some degree of independence from each other, there may not be an overarching institutional message to which analysts must adhere. In that case, the conflicting accounts about PISA may simply be representing conflicting views within the think tank, or the proverbial right hand not knowing what the left is writing.

But the contradiction also speaks to the agency and interests of think tanks. Think tanks want to produce high quality research, but they also aim to find traction in the wider educational debates of the day. Developing a timely, strong narrative to engage policymakers is one course of action through which think tanks pursue their goals. Crisis talk and reference societies are two discursive strategies that make for a compelling narrative. In the next section, I look more closely at the think tank use of reference societies, particularly Finland and China, and how the use of crisis talk puts the think tank literature in the spotlight.

Think Tank Use of Reference Societies

Figure 4.7 illustrates think tank references to China and Finland over time. It confirms 2009 as the pivotal year and charts the subsequent rise in references about China and Finland by both think tanks. Approximately one-third of the texts reference China in some way, slightly less than the number of references to Finland. The China narrative drives the whole discourse, even as the number of references to China itself is a bit below those to Finland.

Figure 4.7: Think tank references to Finland and to China



Fordham uses reference societies more frequently in its work than Brookings does. The references it makes use more colorful and forceful language, promoting the ideas of competition between countries and fear about China's ascendance. Rich (2004) and McDonald (2014) have shown that in general conservative think tanks receive more coverage in the press than do liberal think tanks, in part because they are unafraid to promote ideologically driven ideas and language, and this is borne out in the more frequent use of references to Finland and China, which capitalize on America's relative competitive position in the world order. This is a mechanism by which the Fordham rhetoric is made more memorable.

While both China and Finland are glorified as models of success, the explanations for the success of each are completely different. Figure 4.8 shows arguments for success by country. According to the think tank literature, China's success is predicated on rote

learning and political will, while Finland’s owes its success to its communitarian culture and commitment to equity and teachers. I discuss these themes in more detail below.

Figure 4.8: Think tank explanations of Finnish and Chinese success

THEME	FORDHAM		BROOKINGS	
	FINLAND	CHINA	FINLAND	CHINA
EQUITY	1	0	1	0
TEACHERS	2	0	1	0
COMPETITION	0	2	0	0
LOW COMPETITION	1	0	1	0
ROTE LEARNING	0	2	0	0
POLITICAL COMMITMENT	1	1	1	0
HOMOGENEOUS SOCIETY	1	0	1	0
PISA ALIGNS TO CURRICULUM	0	0	1	0
CHINA'S SCORES UNRELIABLE	0	0	0	1

References to Chinese success. Of the references to China, about half discuss the idea of Chinese success in some detail. Chinese success is explained by a set of factors totally different from those of Finland. It is also explained differently by each think tank. Brookings tries to debunk the idea of Chinese success from the get-go, arguing that “no one really knows how [China] perform[s] on international assessments” (Loveless, 2010, 12). Since only students from Shanghai participated on PISA, the Chinese sample is not representative and therefore results for China remain unknown. As Loveless writes:

“For centuries, Shanghai has been the jewel of Chinese schooling, far ahead of its urban peers and light-years ahead of rural schools. Shanghai’s municipal website

reports that 83.8 percent of high school graduates enter college; the national figure is 24.0 percent” (Loveless, 2010, 12)

The Fordham literature takes a different tack. It describes a tightly controlled education system with a relentless focus on succeeding on high stakes tests. A Fordham blog post provides a laundry list of important factors for success:

“Long school days, really long...No varsity sports. With a school day that long, who has time for a three-hour basketball practice after classes?...a national curriculum... high-stakes testing on steroids... Few questions. Students are generally not permitted to ask questions in class...teachers generally focus on rote memorization of various facts and formulas.” (Joch, 2010)

Each of these reasons also involves a deep political commitment to education. This political will is the crux of the argument in another Fordham discussion of Chinese success (which also ran as an op-ed in *The Wall Street Journal*)

“To be sure, it was only Shanghai (that participated in 2009)...But Americans would be making a big mistake to suppose that this Shanghai results is some sort of aberration...

“How did Shanghai accomplish this?...the city’s near universal education system, its competitiveness (measured by students’ admission to universities and to the best secondary schools) a very high level of student engagement, a modernized assessment system, an ambitious curriculum, and a program of intervening in weak schools.”

“Right now most cities and towns in China don’t have these resources. But tomorrow is apt to be a very different story.” (Finn, 2012, A21)

Finn concludes that “China is bent on surpassing us” and this is the key driver of its PISA success. Educational success has become a competition, and the US will lose unless it gets serious like the Chinese are. However by attributing Chinese success to political will and systemic variables, Finn also makes the Chinese model one that would be exceedingly hard for America to draw on. This is reflected in the dearth of calls for policy borrowing in the think tank literature, which I discuss later.

The references to China also demonstrate a contradictory embrace and rejection of the Chinese model. As the think tanks are publishing literature trying to debunk the idea of Chinese success, they are also publishing literature that glorifies China as a reference society. China’s PISA status helps legitimate it as an international leader, both educationally and otherwise. This is most prevalent when the research is about countries in the developing world. Brookings research about Latin America, for instance,

frequently looks to China as a model of educational success. China is held up as having a low percentage of students performing at the lowest levels on PISA 2009 math; as having a large number of top-ranked universities; for its high number of patents-per-inhabitant; and for its growing economy (Puryear and Goodspeed, 2011; Cardenas, 2011).

Ironically, these reports are published at the same time as the Brookings 2010 education report that argues forcefully that the Chinese sampling was not representative and not valid for comparison (Loveless, 2010).

References to Finnish success. Finland is used as a reference society in about half of the think tank articles (nine out of twenty). The references are mostly simple references, describing Finland as a “top scorer” or a “successful” system. Teachers are described as a key factor in Finland’s educational success: their rigorous training, content expertise, meaningful professional development and autonomy all support high performance on PISA (Puryear and Goodspeed, 2012; Porter-Magee, 2012; Wagner et al., 2012).

Almost a third of the Finland references discuss the idea of Finnish success in detail, but do so in order to debunk the idea. Like the critique of PISA, the critique of Finnish success can be found in literature from both Fordham and Brookings but each think tank focuses on a different aspect of the Finnish success argument. Brookings debunks the notion of Finnish success by arguing that Finland is a top scorer only on PISA, and in fact had slipped to being fairly average when it participated in TIMSS in the late 1990s. Also,

Finnish education reforms of the past two decades may be especially compatible with the content of the PISA exams, so the Finns may be more prepared to do well on PISA than students in other participating countries (Loveless, 2010, 10).

Fordham makes a strongly culturalist argument against the idea of Finnish success. On this view, what accounts for Finland's success on PISA are characteristics that are specific to Finland. These analyses credit "a unique set of value propositions that favored the overall welfare of the group" (Porter-Magee, 2012), and a combination of Finnish history and demographics that allows Finnish students to be successful. These culturalist arguments are double edged. While they acknowledge those specific factors that contribute to success, they also argue that the very factors responsible for success are neither able nor desirable to be transferred to other contexts; the "yes, but..." construction. This creates a paradox in the rhetoric, which celebrates the very factors that cannot be borrowed or lent.

Minimal Calls for Lesson Drawing

This culturalist argumentation does not lend itself to calls for policy borrowing. In fact, the think tank literature, like Allen Luke's article from AERJ, concludes that policy borrowing is not a good option in addressing America's low PISA scores. As one analyst writes:

“The closer you look, the more you realize Finland’s approach works not because it is a universal template of success but, instead, because it was a Finnish solution to which they committed. Americans shouldn’t be looking to slavishly copy these exact hard choices...” (Porter-Magee, 2012)

Such characteristics as having no elite colleges (because higher education is completely state-run), having few Finnish-language learners in classrooms, having lower levels of economic inequality all support Finland’s educational success but are irrelevant to the US situation (“A certain Finnish *je ne sais quoi*,” 2008). Not only is the literature trying to debunk the notion of Finnish success, but it is also challenging the whole notion of policy borrowing and lending, arguing that only a specific constellation of political, cultural and historical contingencies has allowed for the Finnish system to develop as it has. Those contingencies are not applicable to US education reform, and the think tanks stop well short of suggesting that the US reformers should be trying to emulate Finnish success.

The tension between decontextualizing countries in order to make them comparable and recontextualizing them in order to understand the factors that influence student performance is rhetorically and practically problematic. Because the discourse is contradictory, on the one hand arguing that all countries are commensurate while at the same time showing that all countries are specific, it is harder to utilize it for policy. Fordham and Brookings both produce strong narratives about PISA, its strengths and its weaknesses, and play a pivotal role in driving the larger discourse in the media, even

providing language for the US president to appropriate. But the inconsistent interpretation of PISA scores means that there is no clear rhetorical strategy to define the problem, invoke winners and losers, establish ownership or suggest clear solutions.

Conclusion

In this section I presented the findings from the think tank literature. In both the Fordham and Brookings literature, PISA performance is understood as a measure of success: high PISA scores represent a strong education system. There is a simultaneous acceptance of and rejection of PISA, as both think tanks offer stinging critiques but nonetheless accept PISA as a standard by which to measure excellence.

Finland is referenced as a success story, but China emerges as the critical element of the overall PISA discourse. Chester Finn's blog post likening China's top scores on PISA 2009 to the Sputnik crisis of the 1950s rearranges the PISA discourse. Think tanks measure their success in no small part by the influence their ideas have on public policy debate. This goal incentivizes think tanks to produce research that is not only high quality, but also timely, relevant and catchy – as with the Sputnik metaphor. Think tanks are directly tied to domestic political action, sometimes influencing the action, sometimes responding to it. With PISA, a narrative that begins as a series of caveats about using and interpreting PISA scores becomes an admonition for the United States to improve, lest it lose its competitiveness on the world stage.

For think tanks, the marketing and dissemination of research with an eye towards the timely release of reports on ‘hot’ topics, is of crucial importance. This interest is evidenced in both the timing and language of the think tank literature, which is written to attract attention and gain traction in the world of elite policy making. The most influential venue for think tank ideas is in the media, and it is the role of the media in the American PISA discourse to which I now turn.

Media

In this section, I examine the American media discourse about PISA using an original dataset of 73 articles published between 2000 and 2012 from four American media sources. The media make up the largest portion of the sample texts. Media not only provide information about PISA to the public but also play a key role in determining who is given voice to publicly shape the PISA discourse. In order to ensure as broad an understanding as possible of how the media engages in and frames educational debate regarding PISA, my sample includes a combination of elite (*The New York Times*, *The Wall Street Journal*, and *The Washington Post*) and popular (*USA Today*) press, and represents media outlets from across the political spectrum. It also includes media that have an online presence as well as a traditional print presence.

I find that the media discourse is the most robust and influential part of the public dialogue about PISA. The data show that the ways in which the press frames PISA scores, as well as to whom the press gives voice in the debates, aligns with dominant

interpretations of US education since the early 1980s. Education is constructed as a function of economic development, and there is cause for concern when the United States is not keeping up with its geopolitical rivals. The US must take seriously calls for education reform or risk falling behind its competitors. However, there is no clear agreement on what path this reform should take. In addition, the press certifies very few voices to participate in the construction of meaning around PISA. Instead, the discourse relies on only a few expert voices that imbue PISA with too many meanings to have much rhetorical bite

The findings are presented in four parts. First, I discuss general patterns of media discourse. Next, I discuss the use of crisis talk and reference societies in the media, showing the influence of think tanks and the way in which the media capitalize on China's top scores. I show how references to China and Finland energize and direct the discourse, but do not result in calls for specific policy change. The third section examines media framing in detail. I discuss the three major frames used by the media: (1) the US education system is in trouble; (2) American economic growth depends on education; and (3) the US system needs reform. Lastly, I address the issue of voice, focusing on the profoundly elite nature of the discourse. I also show how elite stakeholders use PISA to problematize American education in vastly different ways.

From “Good News, Bad News” to “Stunned Educators”: Two Phases of US Discourse about PISA

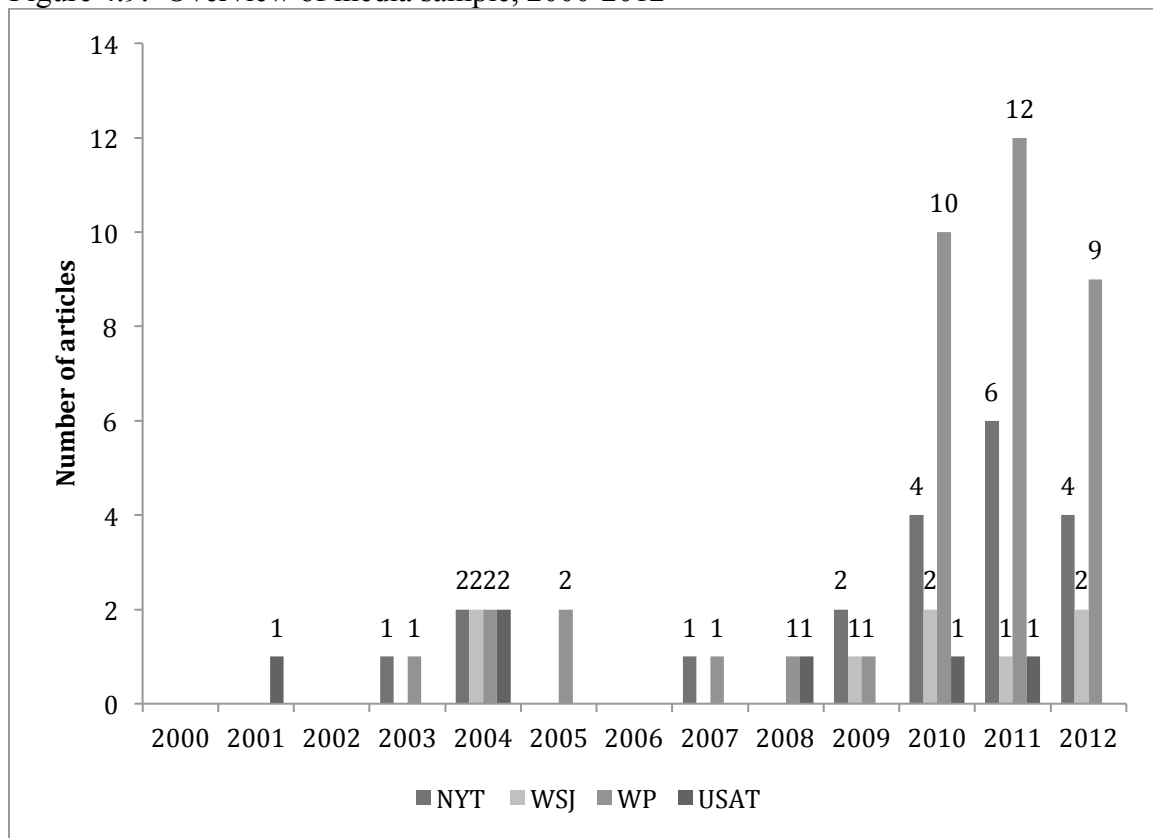
During its first decade, the status of PISA in the United States changed dramatically, from an unknown testing instrument to a bellwether of educational excellence. That this transformation occurred in a mere four rounds of testing reflects not only the rhetorical power of rank and competition, but also the power of the media itself in shaping political discourse. Within the media, PISA is understood primarily in two ways. First, the media emphasizes that PISA is a novel and model assessment instrument, based on applied knowledge rather than curricular recall. Second, like the think tank literature, PISA is used as a measure of national educational success. In the media, if a country performs ‘well’ on PISA, it is understood to have a ‘successful’ education system.

Figure 4.9 shows the dramatic increase in media coverage by year throughout the first decade of testing. While the media coverage grew over time across all four media sources, the sample is better understood if broken into two distinct periods. Phase I covers the administration and release of results for the first three cycles of PISA, from 2000-2009. Phase II covers the administration and response to the fourth cycle, from 2010 through 2012. As I explore in detail later in this section, the explosive growth in PISA coverage is explained by Chinese participation on PISA. Prior to the 2009 cycle, there is minimal coverage of PISA in the US press. PISA is a new assessment instrument, and the top scoring country is Finland, about which most Americans know little. The coverage remains fairly flat until the PISA 2009 results are released, and

China is revealed to be the top scorer. The political economy of the discourse shifts completely and, as in the think tank literature, coverage of PISA skyrockets.

Phase I: 2000-2009. The first phase, from 2000-2009, saw fairly minimal media coverage, with an average of 2.3 articles a year published across all sources. Some years, as in the first round of PISA in 2000, received barely any media coverage at all. The only article to cover PISA 2000 is found – somewhat ironically, given the highly elite nature of the overall discourse -- in *USA Today*, the most populist source in the sample and the one that has published fewest articles about PISA overall.

Figure 4.9: Overview of media sample, 2000-2012



The initial piece is published as a news item in December 2001, on the release of the first cycle of test scores. The short article leads with a focus on the *success* of American students on PISA, saying, “About one in eight American 15-year olds have reading skills that rank them among the top 10% of students in industrialized nations” (Henry, 2001, D8). The article describes PISA as a new, “unique” international study, though it does not explain why. A few key findings – including the US position relative to a handful of other countries, including Finland, Korea, Japan and Brazil (as the low scorer) -- are put into bullet points, but with no analysis. As is typical of *USA Today*, the article includes a simple graphic, which ranks the 32 countries that participated.

Though the article acknowledges an achievement gap -- compared with other developed countries, “US students are just average in reading skills... mainly because of the large percentage of poor readers” -- it makes no detailed arguments about equity or closing the achievement gap. A quote at the end of the article foreshadows the argument linking PISA to economic growth that will become prevalent in later years. An official at the US Department of Education’s National Center for Education Statistics (NCES), notes that US students will need to compete for jobs in the global marketplace, and the article describes the results as a “possible wake up call for America’s youth”⁴⁵ but this hint at competition is not a focus of the article and the link is lukewarm at best. The then-deputy

⁴⁵ It is interesting that the author frames this as a wake up call to students themselves, suggesting that they -- rather than teachers, curricula or other systemic factors -- are responsible for their performance. This reflects the same neoliberal ideology behind many education reform initiatives of the 1990s and 2000s, in which there is a shrinking role for the state and a responsibility for schooling is that of parents and students.

director for education at the OECD concurs, remarking, “the best students in the United States perform as well as those anywhere” (Henry, 2001, D8).

In this first look at PISA in 2001, US education is described fairly positively. The results are straightforward, and the article is more press release than analysis. The US is performing at an average level but is characterized as doing fine. Other countries are neither glorified nor vilified. There are no calls for policy reform, or policy borrowing. The PISA assessment itself is new, so basic explanation about the structure of the test is provided. The coverage is simple, short, and newsy, and the article is brief.

Between 2003-2005, media coverage increases slightly, to an average of four articles per year across all media sources. There is a spike in coverage in 2004 when, due to a scheduling coincidence, PISA 2003 and TIMSS 2003 results are released simultaneously. Eight articles are published that year, several of which discuss US performance in more detail than earlier accounts. PISA is becoming better known and is being touted as a model assessment, as it “goes far beyond the multiple choice questions many students see on [typical] standardized tests” (Toppo, 2004, D7). There is no objection to the fact that PISA is not aligned to curricula or that it has a particular view on knowledge that some will later charge is ideologically biased (Loveless, 2011). Instead, the assessment is described as “uniform” and valid, with “OECD monitors ensuring it wasn’t selectively given” to students but rather was given to statistically representative samples (Kronholz, 2004, B1).

This early coverage takes a ‘good news, bad news’ approach. The good news, coming from TIMSS, is that “American eighth-grade students have made significant gains in math and science compared with their worldwide counterparts”. However, the bad news comes from PISA, where “things begin to look a little bleak” (Dobbs, 2004, A02) when students are asked to apply those skills to real-world situations. The negative coverage paints US students lagging behind their peers in other countries, and then-Secretary of Education Rod Paige calls the 2003 results a “blinking warning light”⁴⁶ (Toppo, 2004, D7).

The negative tone continues throughout Phase I, and much more attention is paid to the ‘worse’ results on PISA rather than the relatively stronger results on TIMSS, particularly with regard to math scores. While the main subject of the PISA 2003 cycle is mathematics, coverage of the 2006 cycle also emphasizes math performance (AP, 2007). Literacy performance gets less attention throughout the discourse. The “hard,” quantifiable skills of math get linked to science and technology, which get linked to competitive labor markets and economic growth, as in this 2004 article:

“The study [PISA 2003]... indicated that huge numbers of US students can rarely do math, meaning the US lacks the advantage of a generally well-educated population, which can...spur growth.” (Kronholz, 2004, B1)

⁴⁶ Paige credits the better results on TIMSS to “a greater emphasis on higher standards in the classroom” at the NCLB, passed under President Bush. See Toppo (2004).

This argument is typically embraced by political conservatives and in Phase I this viewpoint is heard most loudly in *The Wall Street Journal*. However, this framing quickly loses any partisan stance after the release of PISA 2009 results, as all media link economic growth to PISA performance as a key element of the crisis narrative.

After 2004, coverage recedes. The release of the 2006 cycle results garners little media attention though the tone remains the same. Though US rank has not gotten statistically worse, the articles still focus on the poor performance of US students. Poor performance is coupled with economic growth more strongly, as it is argued that American students are now competing with “not just Brooklyn and Boise anymore. American students are competing with Bangalore and Beijing” and US rank is described as “desultory” (Lynch, 2008, B1). The fact that PISA is not tied to curricular knowledge is glorified as a model “which should particularly interest employers” (Kronholz, 2004, B1), since it “insists on the functional ability to extract knowledge and apply it to personal, social or global situations” (Pfeiffer, 2009).

Phase II: 2010-2012. If the first phase is mostly innocuous coverage of PISA, the second phase demonstrates how quickly and powerfully media can define education policy issues. The second phase of coverage is from 2010-2012 after the release of the 2009 results. There is a spectacular shift in public discourse, as the quantity of articles, the tone of the articles and the placement of the articles changes dramatically. Phase II also solidifies the extent to which PISA becomes taken-for-granted in the discourse. In

analyzing the media representation from this period, the data show that PISA was both accepted and critiqued (albeit minimally), and appropriated as evidence in all manner of education policy issues almost overnight.

Breadth and depth of media coverage. There is a dramatic increase in the sheer number of articles referencing PISA in the American press from 2010-2012 (see Figure 4.10). Fifty-three articles, or 72% of the sample, were published during Phase Two, an average of more than 17 articles per year (up from 2.3 articles per year in Phase I). There is also a change in the depth with which the press covers PISA. In Phase II, nine articles discuss PISA in depth, which is more articles than had been published about PISA in depth during the entire previous decade.

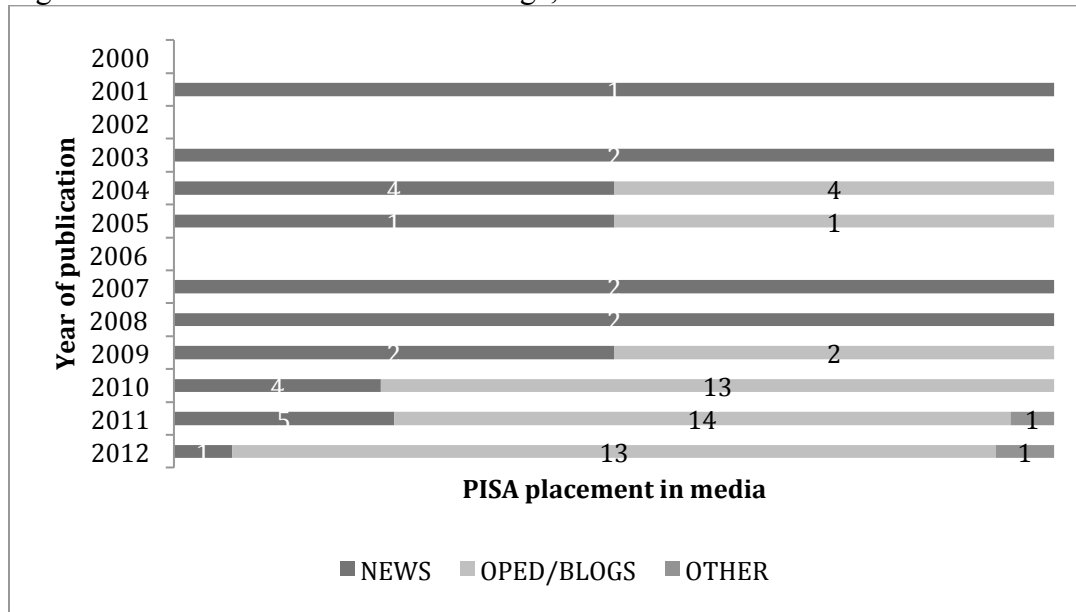
PISA is more frequently mentioned in left leaning and policy oriented media than in the conservative and popular press. Slightly more than three-quarters of the articles are found in either *The New York Times* or *The Washington Post*, with the remaining quarter of articles split between *The Wall Street Journal* and *USA Today*. *The Washington Post* records the most articles about PISA, largely because of Valerie Strauss's popular education policy blog, where posts about PISA exploded after the 2009 results were released.

Placement of media coverage. The expansion of public discourse about PISA is reflected not only in the aggregate number of articles published and the depth in which

PISA is discussed, but also in the placement and types of articles published. The placement and type of article are important provide insight into how rapidly PISA has moved from being considered a news headline to being used as a more integrated, taken-for-granted measure of educational excellence.

Figure 4.10 identifies coverage by whether it was news, opinion (op-eds/blog posts) or other coverage (e.g. feature, business). In Phase I, articles were published primarily in the national news sections of the print version of the paper. Of the seventeen articles published through 2008, 70 % were from the news section, and the remaining 30% were in the form of editorials, in direct response to the 2003 release of results. There were no blog posts mentioning PISA in any of the media outlets. During Phase II, after the release of the 2009 results, coverage in the news section continued while editorial and blog placement skyrocketed. The continuous increase in coverage of PISA suggests that not only are PISA results seen as increasingly newsworthy, but also that PISA is increasingly accepted as a measure of educational excellence. PISA moves from being a news sensation at the triennial release of results, to a topic that is deeply integrated into public discourse where it is used as evidence in ongoing educational debate.

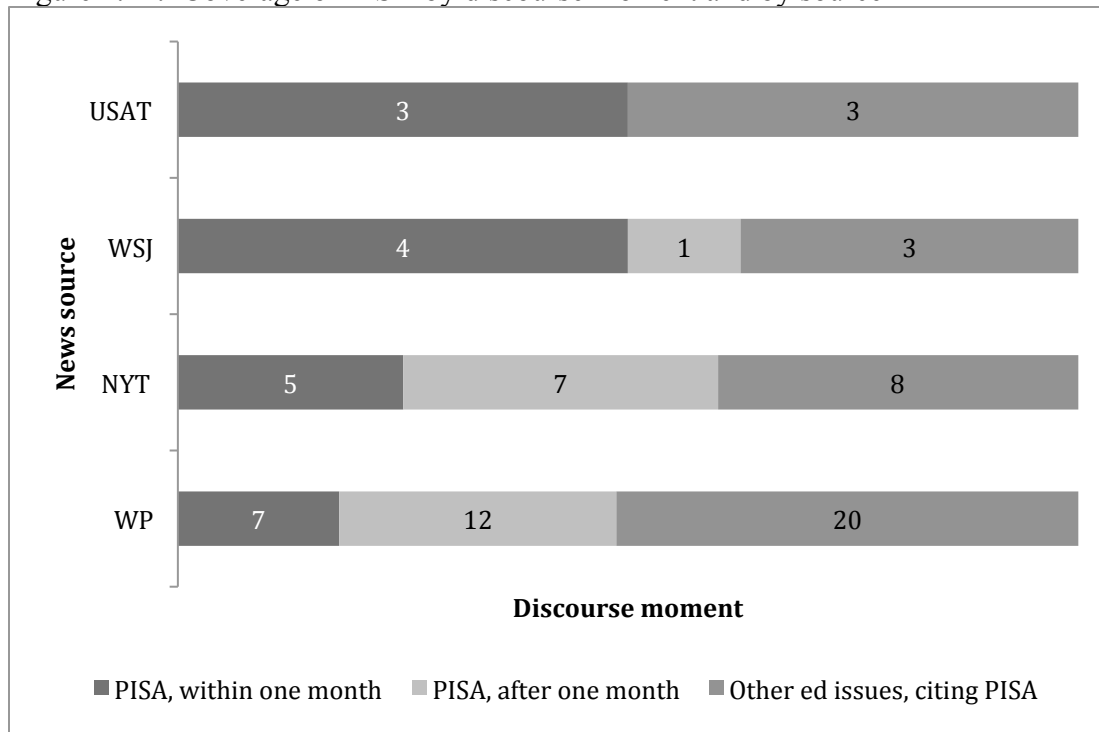
Figure 4.10: Placement of PISA coverage, 2000-2012



Timing of media coverage. The data also show a shift in the timing of articles about PISA. Figure 4.11 presents data on the context of article publication. I distinguish between three discourse moments (following Ferree, et. al., 2002) in which the articles appear: (1) those articles that are published within one month of the release of PISA data and treat the release of results as news; (2) those that are about PISA results, but are published more than one month after the release date; and (3) those articles that reference or cite PISA in the context of other educational or social issues. All four media outlets covered the release of results as a news event consistently during the twelve years covered in the sample. But all papers, especially *The Washington Post* and *The New York Times*, also incorporated PISA into coverage of other education policy issues. This

reflects the ‘taken-for-granted’ status that PISA achieves in the public discourse within only a couple of years.

Figure 4.11: Coverage of PISA by discourse moment and by source



Minimal PISA critique in the media. Throughout both Phase I and Phase II, there is little critique of PISA, either on technical grounds or by whether such scores are an appropriate measure of national education system success. A handful of blog posts critique PISA data, noting that comparing average national level scores is an oversimplification that does not capture the whole of an education system’s success. Comparing national level scores says nothing about students’ levels of confidence, enjoyment or valuing education (Zhao, 2012); nothing about the socio-economic conditions that contribute to student performance (Porter, 2012; Hatch 2010); and nothing about how creative or entrepreneurial students are (Sener, 2010; Zhao, 2012). If rank is

compiled using different measurement data, country ranking can change dramatically (Matthews, 2011). Dillon (2010) also raises the objection that the 2009 China sample may not have been representative. Clearly, though, scoring well on PISA means the education system is strong. This argument is made more vivid by the incorporation of direct comparisons of the US to various high-achieving countries. I turn to externalization to specific foreign examples in the next section.

Media Use of Reference Societies and Crisis Talk

Phase II is perhaps most notably characterized by a narrative of decline, one of the most powerful kinds of political storytelling (Stone, 2001). If the tone of Phase I was that of a depressed, cautionary tale, the tone of Phase II articles is aggressive, sometimes verging on hysterical. Shanghai, representing China, participates in PISA 2009 for the first time and is ranked first in all three subject categories. As discussed in the previous section on think tanks, the Fordham Institute's president Chester Finn's forceful op-ed is printed in the Wall Street Journal. It begins:

“Fifty-three years after Sputnik caused an earthquake in American education by giving us reason to believe that the Soviet Union had surpassed us, China has delivered another shock...Shanghai's teenagers topped every other jurisdiction in all three subjects...”

And continues later:

“...until this week we could at least pretend that China wasn’t one of those countries that was a threat... We could allow ourselves to believe that China was only interested in building dams, buying our bonds, making fake Prada bags... But we could comfort ourselves that their curriculum emphasized discipline and rote learning, not analysis or creativity.

Today that comfort has been stripped away. We must face the fact the China is bent on surpassing us, and everyone else, in education.”

(Finn, 2010, A21)

As I discussed previously, the likening of China’s PISA success to the launch of Sputnik frames the issue as a crisis of potentially serious proportions. President Obama picks up on the analogy and opines that, “Fifty years later, our generation’s Sputnik moment is back” (Dillon, 2010, A22). The symbolism of PISA performance crosses party lines, as the Democratic president appropriates the Sputnik description from a leading conservative pundit. The urgency of the rhetoric contrasts with the language of the first phase or articles, which was concerned but not hysterical. US Secretary of Education Arne Duncan echoes the President in his comments on the release of PISA scores:

“We have to see this as a wake up call... I know skeptics will want to argue with the results, but we consider them to be accurate and reliable, and we have to see

them as a challenge...we can quibble, or we can face the brutal truth that we're being out-educated." (Dillon, 2010, A22)

Duncan's 'wake up call' was widely re-quoted throughout the blogosphere and in multiple news articles.⁴⁷ China's scores become evidence of America's decline, and PISA performance suggests that the US is in a dangerous position because of its PISA rank.

Phase I references to Finland and China. Figure 4.12 shows the number of references to Finland and China from 2000-2012. The use of references in Phase I is minimal; there are few in-depth references to Finland and none specifically to China,⁴⁸ since Shanghai had yet to participate. When Finland is referenced, it is often along with other top-ranked Asian countries (e.g., Japan, South Korea, Singapore) but without any discussion or possible explanation of those countries' successes. Instead, all of the mentions are simple, significantly lower than the 20% decision rule used to categorize a reference as simple or in-depth.

There are only two articles in Phase I, from 2005, which discuss Finland in depth. As associate editor for *The Washington Post* spent three weeks in Finland and published two features about its educational system. The articles praise Finland for its success, and the

⁴⁷ For instance, Duncan's response was requoted by the New York times in Guttenplan, NYT, 12/8/10

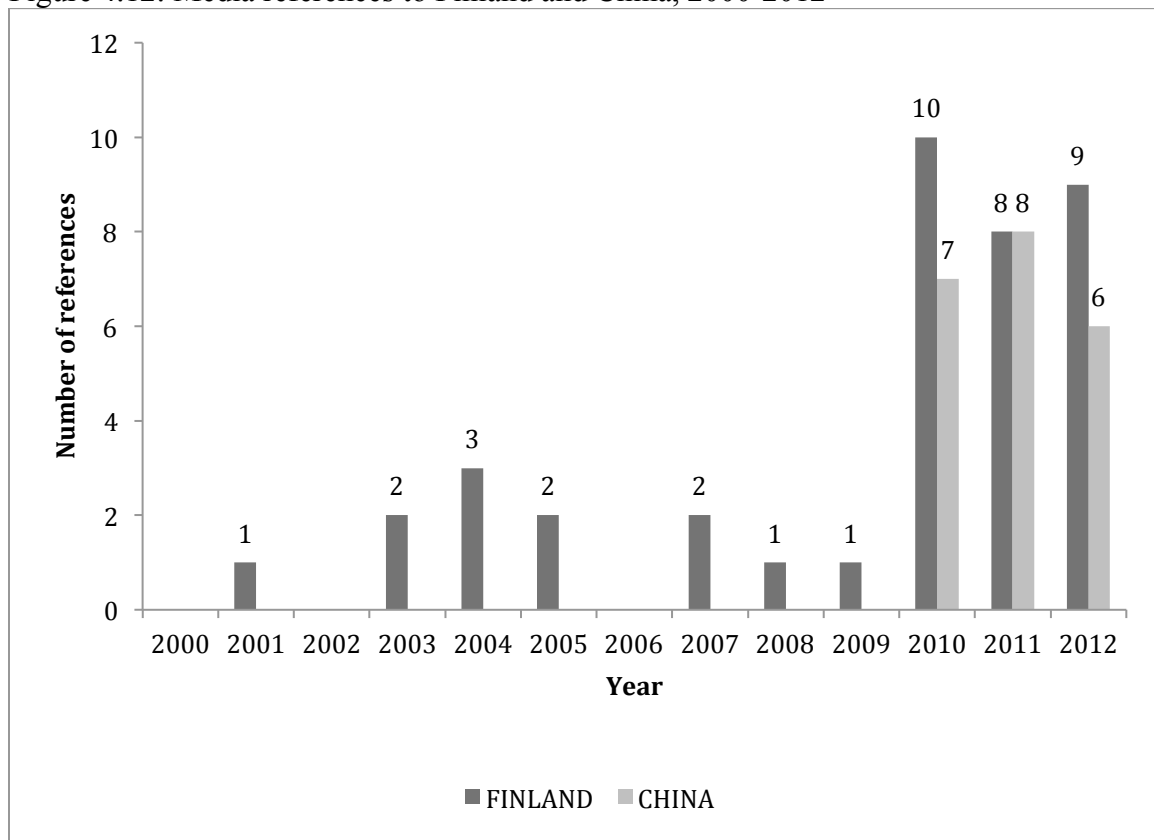
⁴⁸ In Phase I there is an additional mention of China in a 2008 USA Today article, though it is not specifically about China's performance on PISA: "It's not just Brooklyn and Boise anymore. (American students) are competing with Bangalore and Beijing," says Bob Wise, president of the non-partisan Alliance for Excellent Education." (Lynch, 2008)

author credits combination of political will, culture, and teacher quality with Finnish success. Ultimately he concludes:

“I found Finnish society beguiling on many levels, but in the end concluded that it could not serve as a blueprint for the United States. National differences matter.”

(Kaiser, 2005, B1)⁴⁹

Figure 4.12: Media references to Finland and China, 2000-2012



⁴⁹ Again capturing the tension between decontextualization and recontextualization, this quote is also reminiscent of the early comparativist Sir Michael Sadler who cautioned that, “We cannot wander at pleasure among the educational systems of the world, like a child strolling through a garden, and pick off a flower from one bush and some leaves from another, and then expect that if we stick what we have gathered into the soil at home, we shall have a living plant. A national system of education is a living thing...it has in it some of the secret workings of national life.” (Sadler, in Bereday, 1964, 310).

Though Kaiser says that Finland should be an inspiration to education reformers, he stops short of pointing to specific policies that might be borrowed. Instead, he argues that, like the Finns, if the US wanted to make long-term systemic changes, it would be possible with the right combination of political will, political culture and popular support.

The use of references to both Finland and China increases dramatically in Phase II.

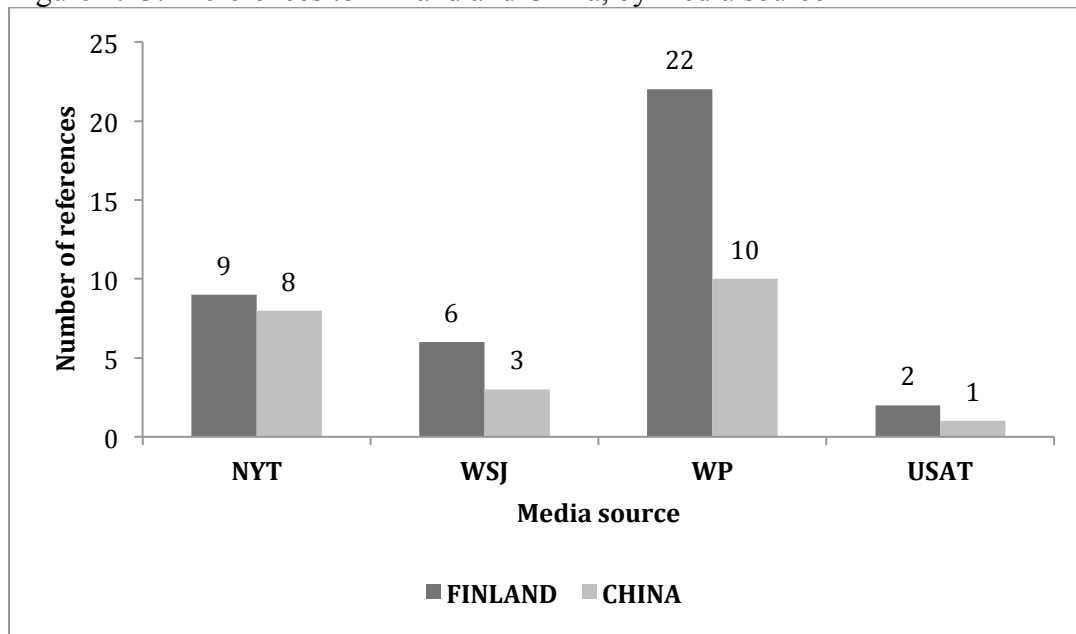
Figure 4.12 shows that not only do references to China explode, but references to Finland do as well. China's top rank ignites the whole discourse to re-focus on comparisons to specific countries, especially the historic top-scorer, Finland along with the new threat, China.

Figure 4.13 shows how each media source uses references to Finland and China. All media use reference both countries as part of their PISA narratives, though USA Today uses this strategy the least. The Washington Post comprises the largest part of the sample (54%) and unsurprisingly has the most references to foreign examples. Most of the Washington Post references to Finland and China come not from news articles but in blog posts, suggesting that references to elsewhere are being used as evidence not only in news stories but throughout the year by individual blog posters, raising the profile and discursive power of PISA.

Phase II references to Finland and China. Looking again at Figures 4.12 and 4.13, the over-time analysis reveals an important pattern. China's success on PISA

rearranged the entire Phase II discourse around references to foreign examples, including those to Finland, even though Finland previously been a consistent top scorer. The release of 2009 results with China as the surprise top scorer resulted not only in the publication of references to China, but in a huge increase in publication of references to Finland. On average, media published 1.2 articles a year with a reference to Finland through 2009, but once China came out on top, references to *both* China and Finland rise. Between 2010 and 2012, media published an average of 14 articles per year referencing Finnish success and more than 7 articles per year referencing China.

Figure 4.13: References to Finland and China, by media source



However the number of articles in which the success is discussed in detail favors China over Finland. Of the articles that mention China, 35% of them discuss Chinese success on PISA in some detail, while only 26% of articles that mention Finland discuss Finnish success in depth. The US being “beaten” by China gets more press attention and

energizes the entire discourse by driving coverage about both reference societies.

Chinese success ups the ante amongst education elites, and the narrative turns to portraying the rise of China at the expense of the US. A *New York Times* column described it:

“It can be win-win...for the world if China is doing better. But not for America...if [we are racing against] a country that is not just out-saving us and out-hustling us, but is also starting to out-educate us.” (Friedman, 2010, A35)

Explanations of Finnish success

The arguments about Finnish and Chinese PISA rank emphasize different explanations of success between the two countries. Figure 4.14 compares what the media say about explanations for success in each country.

Figure 4.14: Media explanations of Finnish and Chinese success

FINNISH SUCCESS	CHINESE SUCCESS
Teachers (28%) <ul style="list-style-type: none"> • Recruitment • Rigorous training • Ongoing professional development 	Education Policy (30%) <ul style="list-style-type: none"> • Teaching to the test • Rote memorization • Want to lead the world
Culture (26%) <ul style="list-style-type: none"> • Communitarian (v. individual) 	Culture (21%) <ul style="list-style-type: none"> • “Asian” model/commitment • Confucian reverence for education
Equity (19%) <ul style="list-style-type: none"> • Inputs • Outcomes • Overall economic equity 	

Teachers. Regarding Finland, media cite ‘teachers’ as the most important explanation of Finnish success, 28 % of the time. This argument captures a range of specific teacher-related features, including teacher recruitment, training, pay and/or ongoing teacher development. Finnish teachers are deemed successful because of highly systemic factors -- Finnish teachers are recruited in a highly competitive process; they undergo rigorous teacher training with an emphasis on deep content expertise; teachers are given serious ongoing teacher development; their pay is good relative to other professions; and they are given a great deal of autonomy in the classroom. This is a potentially potent argument for reformers because it is possible to replicate these factors in other systems.⁵⁰

In spite of this, there are no consistent calls for specifically borrowing from the Finnish system. Instead, a typical explanation includes reverence for Finnish teachers and the respect they receive in Finnish society: “The key.... it’s the people. The high quality of Finnish education depends on the high quality of Finnish teachers” (Kaiser, 2005, B1). Teacher education programs are “highly competitive” and “rigorous,” and principals and parents alike trust teachers (Ravitch, 2011). All detailed articles about Finnish success reference teachers. Finnish experts are quoted as often as American commentators, but with the caveat that their success is specific to Finland. Again, there is a tension between commensuration and decontextualization.

⁵⁰ This is not to suggest that systemic reform is easy. It is of course highly contested and political. But all things being equal, calls to borrow policy are may be more concrete than calls for cultural or political changes.

Culture. Teacher explanations for Finnish success are closely followed by cultural explanations, referenced 26 % of the time. References to culture refer to the Finnish orientation of putting community interests before that of the individual. Whereas in the United States education is often seen as a “private effort leading to a public good... in Finland, education is viewed as a public effort for a public purpose,” (Sahlberg, 2012) where the key beliefs are “collaboration, trust, responsibility and autonomy” (Ravitch, 2010) not competition.

Equity. The cultural explanations dovetail with the third explanation for success, which is equity. In 19 % of the articles, equity is cited as the root of Finnish educational success. Equity in Finland comes in two types. Equity of inputs – equal allocation of resources to schools, a strong early childhood welfare system, and not only of resources to schools but even more importantly an equality of opportunity for every child to succeed no matter what school he or she attends – and equity of outputs, the fact that the variation in between school performance is quite low, meaning that children get a similar education no matter what school they attend.

Explanations of Chinese Success

China, as an economic and geopolitical rival of the United States, is the catalyst that drives American PISA discourse. But there are no calls for lesson drawing from China. The media explain Chinese success by factors that are not associated with the US system and therefore offer little leverage as a model for emulation.

Teaching to the test. The reasons media employ when explaining Chinese success run the gamut. Chinese education policy is credited with China's success on PISA, with 30 of the articles referencing this explanation. However, the policy that is credited with Chinese success is that of 'teaching to the test.' On this view, the Chinese succeed because rote memorization is encouraged at the expense of creativity, and students spend excessive time in school preparing for the test.⁵¹ This is an especially interesting argument since PISA is not tied to any school-based curricula, and teaching to the test should be moot.

Culture. This argument dovetails with the next most frequently cited reason for Chinese success, Chinese culture. Twenty-one percent of references cite cultural explanations. Here, the media highlight a particularly "Asian" commitment to education linked with a Confucian history in which teachers and education are revered (Lynch, 2008; Chua, 2011; Kristof, 2011). The reliance on cultural arguments is unexpected since these are arguments that place success at the mercy of a confluence of political, historical and social factors that are time and context specific. Culture is rarely transferrable. It is bounded by the political and social factors of a particular time and place. The question is why the narrative employs cultural explanations so frequently, if they have little to offer in terms of lesson drawing.

⁵¹ Some articles also mention the excessive out of school time spent in private tutoring, for example, Zhao (2010).

It may be precisely for this reason that the media generally and elite policy actors specifically use the culture argument. It allows politicians and policy makers a way to acknowledge that the United States scores much lower than one of its economic and geopolitical rivals, but it provides little in the way of concrete problem definition that could be used as a basis for reform. Culture is useful in that gives actors a point of comparison, but is a slippery enough concept that it offers little in the way of policy prescription. The value of the reference is in the rhetoric rather than the substance, and allows actors to stake out their opposition to the state of American education without committing to any viable solution.

Strikingly, none of this rhetoric seems to have presaged any further arguments for policy change. References to Finland, China or PISA in general are used as ‘one more reason’ that the US education system is failing and needs reform, but the specifics of the argument – that certain policies regarding systems, teachers and students should be adopted – never takes hold.

Framing PISA in US Education Discourse

Framing refers to the words, images, style and themes a speaker (an individual, or a media outlet) uses when relating information about an issue to an audience. Issue framing is a key element of political discourse and policy making, as it has been shown to have an impact on attitudes and policy preferences amongst voters, politicians and journalists (Chong and Druckman, 2007). As Benford and Snow (2000) explain, the

concept of framing is analytically useful in illuminating how ideas are generated, diffused and mobilized. The approach views actors as signifying agents who are actively engaged in constructing meaning about social ideas. In doing so, they interpret phenomena and create shared meaning, privileging certain interpretations over others thereby constraining other problem definitions and solutions.⁵²

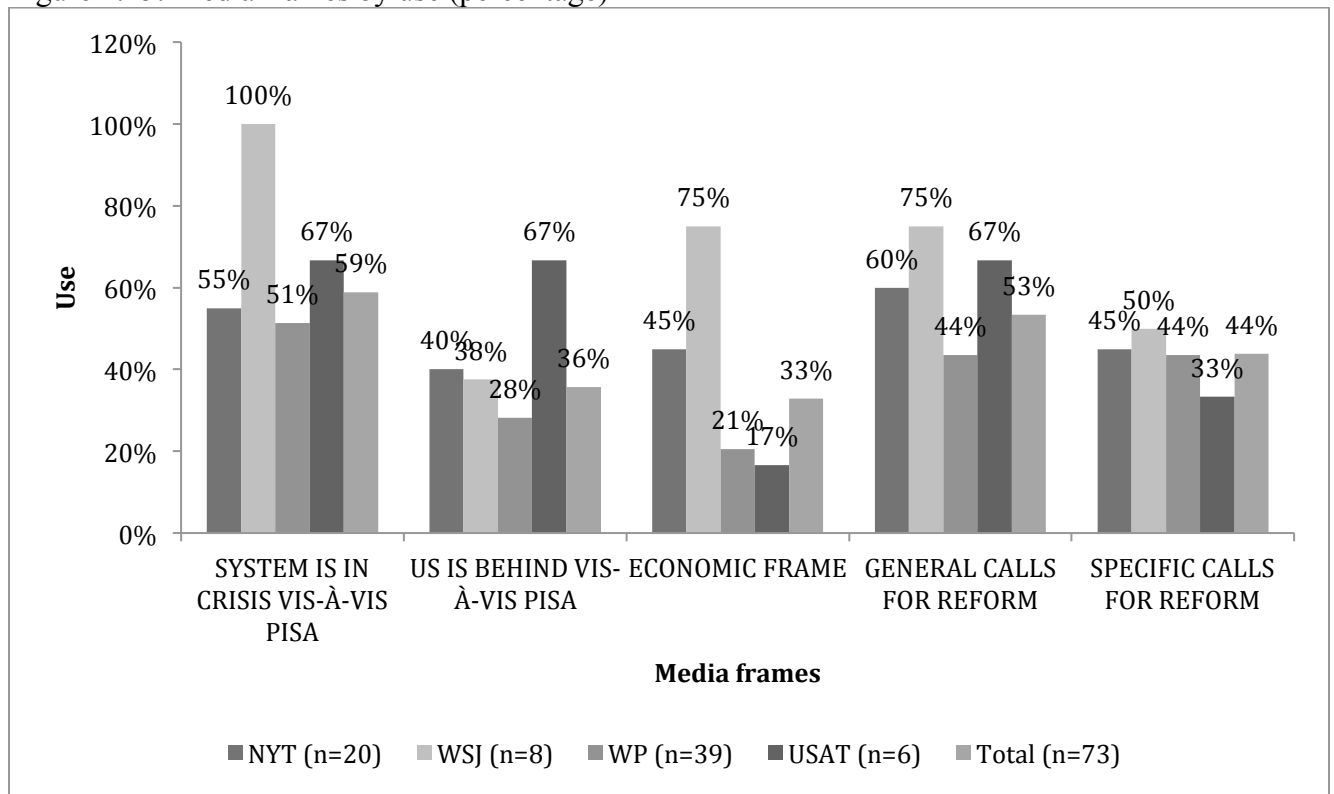
One important element of framing concerns resonance, the degree to which choices about framing are more or less effective because they align with pre-existing cultural conditions and paradigms (Benford and Snow, 2000; Mehta, 2011). The frames in the PISA discourse resonate because they are culturally consistent with public understandings of education and the views of institutional elites (McAdam, in Benford and Snow, 2000, 626). PISA rapidly becomes integrated into the larger educational discourse because it aligns with understandings of the US education system that have been the zeitgeist for several decades. The media frames are basically uncontested. The three major discursive frames are not mutually exclusive and many articles employ multiple frames to make their arguments.

Figure 4.15 offers more nuance to the major frames by introducing sub-categories. The first frame, that the US system is in trouble, has two sub-categories: (a) that PISA scores show the US education system is in crisis, and (b) that PISA scores show the system to be

⁵² Political sociology focuses on framing as a function of collective action (see, for example, Benford and Snow, 2000) and political mobilization, while political science addresses the issue as a function of the policy process (see Stone, 2001; Chong and Druckman, 2007; Schmidt, 2008.). I draw on both literatures in my analysis.

behind, but not in a full-blown crisis. The second frame is an economic one, presented as one category, though I discuss further nuance below. The third frame, that the US needs education reform, is also divided into two categories: (a) general calls for reform, and (b) specific calls for reform. Figure 4.15 also shows which sources employ which frames most often. The frames are not competitive or particularly partisan and there is remarkable agreement across sources on frame usage. This consensus makes for a strong, shared storyline across sources and over time.

Figure 4.15: Media frames by use (percentage)



Frame 1: The US education system is in trouble. The first frame uses PISA scores as a representation of the status of the American education system. All four media

outlets use PISA scores to comment on the state of US education, and this frame is present in more than a third of the media articles. The fact that the United States is not a top-scoring country is symbolic of a variety of ills in the system. This frame aligns with the way in which American education has been represented in the media since 1983, as a troubled, underperforming system that is putting American interests “at risk.”⁵³

Within the status frame, “crisis talk” is a notable sub-frame. It describes the US system as in various stages of crisis vis-à-vis its PISA scores and China. Almost 60% of the status articles describe US education as being in some state of crisis. This is a particularly strong frame in *The Wall Street Journal*, where 7 out of 9 articles portray the US as in dire straits with regard to its PISA performance. As discussed previously, the strongest framing of this crisis is the op-ed piece by Chester E. Finn, who likens the US being outscored by Shanghai as another “Sputnik” (Finn, 2010, A21). But another earlier article describes the “enormous” achievement gap between white and minority students in the US as well as the gap between the US and other countries as “really big deal” (Kronholz, 2004, B1).

While this might initially be viewed as partisan posturing, as an indictment of the Democratic Obama presidency and its education policy, the left-leaning *New York Times* also uses alarmist, crisis language. The crisis talk crosses party lines. *The Times*

⁵³ Phi Delta Kappan/Gallup poll results since 1974 show that Americans, across all years, award “the public schools in the nation as a whole” a grade of “C”, “D,” or “Fail.” The percentage of respondents giving the US a failing grade has increased dramatically since 2005. See Morgan and Taylor-Poppe (2012, 263, Figure 1).

announces its coverage with the headline “Top Test Scores from Shanghai Stun Educators.” Only *The Washington Post* offers a more nuanced approach, by acknowledging the US scores but immediately offering a counter-argument. This follows the trend of the past few decades, in which Democrats and Republicans are more likely to cross party lines to use similar language and ideas in their education policy positions.

The other subframe concurs that the US is not doing well, but takes a slightly softer stance than full-blown crisis talk. These articles report a generalized malaise of the system as evidenced by its achievement gaps, and by a sense that the US is simply falling behind. But rather than seeing poor performance as a full-blown crisis, these articles matter-of-factly note that poor performance on PISA means the US is not ‘competitive.’ Often, this notion of competition meant as specific to economic growth, but this frame encompasses a generalized notion of competition as well. Middling PISA scores support a notion of the US education system being behind, though the articles provide no interpretation of what the test actually tests, or how the scores actually measure lagging American competitiveness.

Frame 2: Education is necessary for economic growth. Economic growth and education have been coupled in the American education discourse since 1983 and the release of *A Nation at Risk*. It has remained a central part of the American education narrative because it has capitalized on dominant understandings about education. Economic output is quantifiable, and therefore privileges data and measurement as policy

evidence. The market mechanisms of economics also align with the neoliberal, and to some extent neoconservative, ideology that has dominated education reform in the US during the past forty years (McDonald, 2014; DeBray, 2006; McGuinn, 2006).

The economic frame is powerfully used in the PISA discourse across all elite media. Almost half of *The New York Times* articles (46%), three quarters (75%) of *The Wall Street Journal* articles, and a quarter of *Washington Post* articles (25%) employ this frame. Only *USA Today* rejects economic approach as the predominant frame, using it in only one article. It is a particularly strong element of the politically conservative arguments, as evidenced by the frequency with which it is used in *Wall Street Journal* coverage.

Education and macroeconomics. The link between education and economic growth occurs at two levels. Of the articles that use the economic growth frame, half define economic growth at the macroeconomic level, where growth concerns national economic development and is measured in terms like gross domestic product (GDP). The press cite several different reports that link education to macroeconomic growth. One, by Eric Hanushek, a well-known Stanford University economist commenting on PISA 2003 scores, says “trailing other OECD countries on education measures may reduce US economic growth by as much as a half percentage point a year” (Kronholz, 2004, B1). Another influential report issued by the consultancy group McKinsey and Company echoes this analysis, reporting that if the US had closed the achievement gap between the US and higher performing countries in the years since *A Nation at Risk*, US

gross domestic product would have been between 9 and 16% higher. The report also estimates that “closing the gap in the US between white students and their black and Latino peers could increase annual GDP by as much as an additional \$525 billion, or about 4%” (Tomsho, 2009). This links a domestic achievement gap defined in racial terms to overall GDP growth.

Education and microeconomics. The individual level returns on education second line of economic argumentation and is included in the other half of the economic frames. In this framing, individual labor market opportunities are constrained or eliminated without a competitive quality education. As one analyst describes:

“The US has rested on its laurels for way too long. The Baby Boomers were the best-educated generation of any in the global workforce. Today’s labor force entrants are not as lucky...” (Lynch, 2008, B1)

Analysts also focus on the racial dimension of the economic achievement gap, noting that how minority students score on the test has ramifications for growth and inequality. On this view, education is imperative for entrance into the middle class for underprivileged minority groups, and addressing the enormous gap “between scores of white and minority groups – who will make up an increasing share of the labor force in coming decades” (Kronholz, 2004, B1) is the key. Or, as one economist puts it, “It’s their (low scoring

minority groups) productivity that will determine economic growth and whether my generation gets Social Security” (Kronholz, 2004, B1).

Dominance of economic reasoning. The economic frame is very strong, not only because it suggests that educational success can be measured in GDP growth and/or employment statistics, but also because there is almost no critique of the link between education and economic growth. While the media publish minimal critique of PISA in general, critiques of an economic rationale for education are especially limited. There is only one article in the sample that offers a rebuttal to the economic argument. Written by economic journalist Robert Samuelson, it notes that, in fact, economic competitiveness depends on more than good schools. Samuelson writes:

“...[schools] are important but not decisive...American schools do as good a job as schools in other wealthy nations. We’re worse than some and better than others. The overall loss of economic competitiveness is likely modest and would be swamped by other factors (government policies, business management, exchange rates, the willingness to take risks).” (Samuelson, 2011, A17)

He also notes that the overall scores do not account for socioeconomic, racial and ethnic heterogeneity that characterize the US. When scores are analyzed by ethnic group across countries by ethnic group “that US schools do about as well as the best systems

elsewhere in educating similar students.”⁵⁴ Nonetheless, this opinion piece is the only voice contesting the link between education and economic development at either the macro- and micro level. The economic frame dominates the discourse virtually uncontested. That this frame fits well with many preexisting education narratives only underscores its staying power. This framing makes education measurable, across a variety of economic measures (e.g., GDP, individual salaries, unemployment rates). It also supports the narrative of decline. It extends a narrative that has been in place for the past thirty years in the United States, that links US economic power to the state of our education system. This is in spite of the fact that the US has been and remains an economic power no matter where it has come out on PISA or other international assessments since 1983.

Frame 3: The US Education System Needs Reform

The third prominent frame is the call for reform of the US education system, mentioned in more than half of the articles (53%), and used consistently across all four media sources. Some articles call for general improvement, without specifying a particular effort on which to focus, as in America “should be raising its standards” (Williams, 2011, A14). But most articles point to closing the achievement gap as the key to improvement. It is important to note the way in which this framing constrains the problem definition. The achievement gap is a function of schooling rather than, for instance, a function of poverty. The frame implicates the education system, rather than the welfare state

⁵⁴ See Feniger and Lefstein (2014) for a systematic approach to this issue. For more discussion about the importance of non-cognitive factors economic growth, see Levin (2012); Heckman and Kautz (2013).

structure, which has obvious implications for responsibility and blame in terms of problems and solutions.

The macrolevel achievement gap. The international/macrolevel version of this argument uses America's mediocre PISA rank as a sign that there is an achievement gap between the US and higher scoring countries. *The Wall Street Journal* focuses on this in an early article: "the age of top-achieving math students in the nation is about half that of other industrialized countries" (Kronholz, 2004, B1). This idea was emphasized in 2009, when the consultancy McKinsey and Company released a report that was covered in the media. The report spoke about the economic implications of the achievement gap, in which it was estimated that:

"If the United States had in recent years closed the gap between its educational achievement levels and those of better-performing nations such as Finland and Korea, GDP in 2008 could have been \$1.3 trillion to \$2.3 trillion higher."
(Rampell, 2009)

The McKinsey report also received coverage in *The Wall Street Journal* and *The Washington Post*. This approach couples the achievement gap frame with economic consequences, but it demonstrates the reliance on rank as a symbol of educational health.

The microlevel achievement gap. Another approach to the achievement gap highlights intranational/microlevel gap between low and high scoring American students. This argument is framed in strong equity terms, often addressing the poverty from which lower-scoring students come. The point is made repeatedly that those students coming from low socioeconomic backgrounds generally do less well than upper income students (Lucido, 2011; Darling-Hammond, 2012; Porter, 2012). Other articles link the domestic achievement gap and economic frames, discussing the gaps as socioeconomic as well as microeconomic issues:

“...greater educational disparity leads to greater income distribution disparity. If we fail to reform our K-12 educational system, we’ll be locking in inequality problems that will plague us for decades if not generations to come...” (Schultz and Hanushek, 2008, A15)

Yet the articles that call on policy makers to address poverty as the root issue of underperformance on PISA fail to provide any specific policy recommendations. There is no discussion of how poverty affects students in material ways or what could be done at the policy level to alleviate these impediments.

Improve teachers. More concrete calls for improvement include those that involve teachers (16% of improvement references). Teacher improvement encompasses references to change teacher recruitment, teacher training, teacher accountability, or

teacher pay. The top recommendation from an OECD sponsored report – entitled “What the US can learn from the world’s most successful school systems” -- is to “make a concerted effort to raise the status of the teaching profession” (Dillon, 2011, A22). Another analyst, remarking on the “sorry state of public education,” as evidenced by 2009 PISA scores, opines that the solution is upgrading the teacher talent pool, by changing the way the US recruits, pays, and holds teachers accountable (Huffman, 2010). Notably, there are no actual teachers who are quoted in these articles. A highly elite group of policy analysts provide the narrative on teaching and learning, rather than anyone in a classroom. Additionally, the teacher narrative exists primarily in the blogosphere, and not as part of the print coverage of the mainstream press.

Improve culture. Another call for improvement is for a change in educational culture. These calls focus on the ideas of trust and respect; particularly with regard to teachers, the arguments are that the low social status and low pay accorded to teachers needs to change. This trust needs to spillover to schools and school leaders as well. There is also a contrast drawn between the general education culture in the US, which emphasizes education as a private good to benefit individual students, and places like Finland and/or China, where education is viewed more as a “public effort serving a public purpose” (Sahlberg, 2012; see also Ravitch, 2011; Kaiser, 2005). In these articles, the call for an improvement of educational culture include more ‘Confucian reverence for education’ (Kristof, 2011), more interventionist government to make sure the education system as well as the economy are running efficiently (Leonhardt, 2011, B1), and an

environment where hardworking students are rewarded and praised by their peers (Matthews, 2009, B03). The calls for reforming education culture on the one hand seem counterintuitive in that culture defies a common meaning and is particularly amorphous to address in policy terms. Yet this argument appears most frequently in *The Washington Post*, geared most directly to policy makers. *The Wall Street Journal* makes no mention of culture as a possible leverage point in the argument for improvement.

Critique of Using PISA as Evidence of a Need for Improvement

Though it is not its own frame, there is a subset of articles that include critiques of the use of PISA as evidence. These articles charge that the PISA results are not representations of the US education system and should be taken with caution. Unsurprisingly, the criticism is most often found in blog posts which function similarly to op-ed pieces as a forum for opinion and criticism. I take up the role of opinion and the blogosphere in the next section, but it is important to note that the criticism is both (a) limited and (b) offered simultaneously to the explosion of attention on PISA in Phase II of the discourse. A few contrarian voices in the discourse question the relevance of PISA scores to US education policy, but those voices do not come through forcefully. Of the 15 articles that offer some point of critique, 14 are individual posts from *The Washington Post's* education blog.

Voice in PISA Discourse

As I suggested previously, in addition to issue framing, the notion of voice is also critical to unpacking how the PISA discourse is constructed. Who is given voice in the discourse reflects who has standing, and whose ideas are considered valid or worthy. Voice and standing are essential components of policy discourse since they determine which actors have legitimacy and power. As Ferree, et al. (2002) note:

“[Standing] refers to gaining the status of a regular media source whose interpretations are directly quoted. Standing is not identical to receiving any sort of coverage or mention in the news; a group may appear when it is described or criticized but still have no opportunity to provide its own interpretation and meaning to the event in which it is involved. Standing refers to a group being treated as an actor with voice, not merely as an object being discussed by others.”
(p. 13)

In order to understand which actors have voice and standing in the discussions about the importance of PISA for US education policy, I coded and analyzed the speech actions of every speaker quoted in the media. Speech actions include only speech that is quoted directly, in quotations, in the articles; paraphrases or restatements of comments from people or reports are not included in the analysis.

The analysis indicates three main findings: (1) There is virtually no public voice in the discourse, despite the fact that education is one of the most public of issues – virtually all citizens have been in an education system at some point.⁵⁵ Teachers, parents and students are almost non-existent in the discussion. Instead, a small, highly elite group of policy analysts and researchers drive the discourse. Related to this is the outsized role the OECD plays in framing the debate in the US; (2) The analysis shows that opinion pieces, rather than news coverage, are framing the discourse; (3) Individual speakers are using PISA as a ‘projection screen’ onto which they make claims that support their particular policy interests. However, there are few calls for actual policy change or concrete policy initiatives.

The PISA discourse is elite. There is extremely low public voice in the discourse about PISA. Those who are closest to the schoolhouse – teachers, parents, and students – have the least say. Elitism is found both in which media sources promote a PISA discourse, and also in which individuals are given voice and standing.

Article level. At the article level, the data show that discourse about PISA is discussed far more often in the elite sources (*New York Times*, *Wall Street Journal*, and

⁵⁵ There has been little work to date to assess whether or not the American public considers American PISA performance a salient issue with regard to education policy. An exception is Morgan and Taylor-Poppe (2012), who test framing effects on public opinion about domestic education spending. They frame survey question in the context of international competitiveness vis-à-vis test scores. They find that 80% of respondents identify China as the largest economic threat to the United States, and almost 50% of respondents feel that the US public education systems is losing ground compared to those countries posing an economic threat. Though Morgan and Taylor-Poppe do not ask specifically about PISA, public opinion may be reflecting the PISA discourse as it has been constructed since 2009.

Washington Post) than in *USA Today*, a more popular press (see Figure 4.10). Just more than half the articles appeared in *The Washington Post* (53%), with 27% appearing in *The New York Times*, 12% in *The Wall Street Journal*, and 8% in *USA Today*. That a majority of references to PISA appear in *The Washington Post*, an elite newspaper geared toward policy makers, supports the idea that PISA has become an important element in policy discourse. The elite nature of the discourse is overwhelming, with 92% of the references to PISA found in one of the three elite sources.

Within the sample there are 112 separate speech acts made by a total of 79 speakers. Sixty of the articles (n=44) include speakers and speech acts. This number should be qualified, however, because it does not include general text from blog posts or op-ed pieces. Arguably such pieces could be considered as speech acts in and of themselves since blogs and op-eds are designed to be vehicles for the authors' opinions. However, I stick to the decision rule that only includes those speech acts that are published in direct quotes as part of the speaker sub-sample.⁵⁶

Speaker level: US government officials. Figure 4.16 provides an overview of the speakers whom are given voice in the PISA discourse. Speakers from elite institutions are privileged in the discourse. Government officials, including the Secretary of Education as well as other officials from policy offices and statistical bureaus, are quoted most often in the discourse. Several senior officials at the Department of Education

⁵⁶ This means some of my findings may be more conservative than if I had included blogs and op-eds as long speech acts in the analysis.

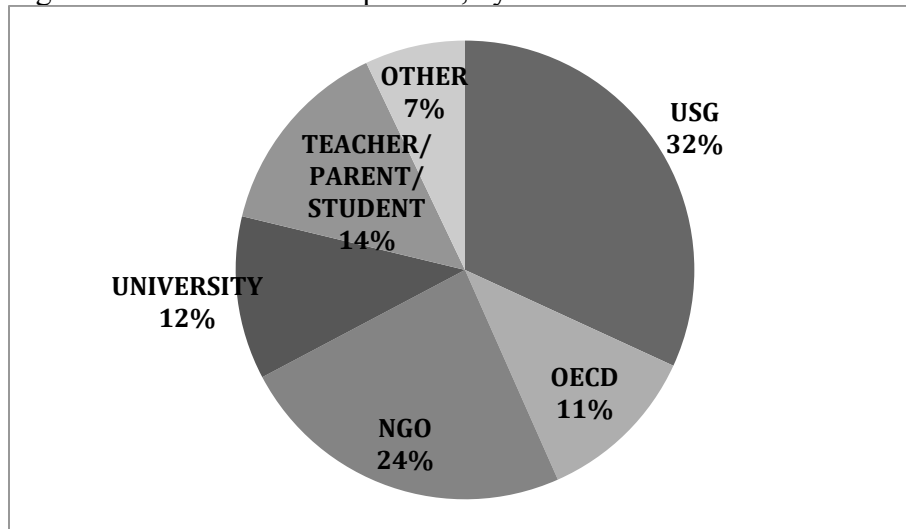
comment over the years on US performance. Quotes from the Secretaries of Education over the years show that their quotes are used to amp up the discourse, signaling the increasing seriousness with which the US government has taken the results. This is most obvious in the reaction to the 2009 results, where Secretary Duncan famously called China's success "a wake-up call" and decried the fact that the US is "being out-educated" (Dillon, 2010, A22).⁵⁷ In earlier cycles, Department staff commented on PISA results. In 2003, Patrick Gonzales of the Department's National Center for Education Statistics (NCES) noted that PISA results showed that US students were not doing as well as their international peers (Arenson, 2004, A29), while in the 2006 cycle, *The New York Times* quoted the Bush administration's staffer at the NCES as simply acknowledging that PISA tests students' abilities to apply "more sophisticated concepts and deeper reasoning skills" (Glod, 2007, A07).

Non-governmental organizations. The next largest affiliation is to non-governmental organizations, almost 24% of speakers. This might initially seem to suggest the discourse would have a strong governmental critique, as many NGOs are mission driven and often take issue with government policy. However, the NGO slice of the pie is far from monolithic, incorporating 22 speakers from 27 organizations. The dilution of voice amongst so many NGO speakers gives NGOs as a category a weaker say in the discourse, as it incorporates disparate ideas from teachers unions, foreign affairs groups, and advocacy organizations.

⁵⁷ In this round of results, the hysteria led to even the US President being quoted, echoing Chester Finn's formulation, saying: "Fifty years later, our generation's Sputnik moment is back.... As it stands right now, America is in danger of falling behind." (Dillon, 2010)

The OECD. The OECD is the third most frequently referenced affiliation and has a disproportionately strong voice in the discourse. In fact, in statistical terms, the data show that the OECD is the only affiliate that is quoted significantly more than other sources ($p=.001$).⁵⁸ Though the OECD Secretary-General, Angel Guriá, and the former head of the Education Directorate, Barry McGaw, were quoted once each about the early rounds of PISA, the voice of the OECD is most directly manifested in the role that Andreas Schleicher has come to play in the discourse. Schleicher, Director of Education and Skills at the OECD and Special Advisor on Education Policy to the Secretary-General, oversees the administration of PISA and the analysis and policy advice published by the Directorate. According to his OECD biography, he also “promotes the work of the Directorate for Education and Skills on a global stage and fosters co-operation both within and outside the OECD.”⁵⁹

Figure 4.16: Overview of speakers, by affiliation



⁵⁸ The only other affiliation that obtains a level of significance in the analysis is “parents,” but this is because parents are quoted so infrequently that one parental quote is significant.

⁵⁹ See <http://www.oecd.org/pisa/contacts/whoswhoinpisa.htm>

Academics. The university is the last category of speaker affiliations. Table 4.17 details the university affiliations within the discourse. White male economists are the individuals most frequently quoted. Here again is the privileging of economy over education, men over women, and the elite university over other sources as the voices of authority in the discourse.

Figure 4.17: Academic speakers and university affiliations

Speaker	Department	Affiliation
Eric Hanushek	Economics	Stanford
Linda Darling-Hammond	Education	Stanford
Lawrence Katz	Economics	Harvard
Richard Murnane*	Economics	Harvard
Richard Freeman	Economics	Harvard
Stephen Walt	Government	Harvard

**While Murnane is the Thompson Professor of Education and Society at the Harvard Graduate School of Education and an economist by training, he is quoted in the paper as a “Harvard economist.”*

To make an additional point about speaker gender in the discourse: though 76 of American teachers are women,⁶⁰ women are only 15% of all of the speakers in the sample. Of those twelve women who are quoted, only one is directly involved in policymaking (former Bush administration Secretary of Education Margaret Spellings), and none is a practicing teacher. None of the op-ed columnists is a woman, though the columnist in charge of the influential *Post* blog is a woman. Of the 112 speech acts, only 15 are quotes from women. As mentioned previously, teachers, parents and students make up only 14% of the speakers.

⁶⁰ U.S. Department of Education, National Center for Education Statistics. (2010). *Teacher Attrition and Mobility: Results from the 2008–09 Teacher Follow-up Survey* (NCES 2010-353) retrieved on September 1, 2014 from <http://nces.ed.gov/fastfacts/display.asp?id=28>.

Others. Additionally, there is an “other” category of speakers. Though this category incorporates only 7% of the speakers, it is telling as to whom the media deems qualified to comment on educational affairs. These “other” voices include a movie star; a technology entrepreneur; a Wall Street hedge fund manager; a journalist and a TV producer; and a public intellectual, amongst others. On the one hand, it substantiates the idea that education is by and of all citizens, so all citizens would have a say. None of these speakers, however, is identified in a role that would directly connect them to the education system (e.g., “former student,” “contributor to local PTA,” etc.). Rather, their professional affiliations *outside* of the education world are what give them credence as sources. James Simons, a “Wall Street investor” and hedge fund manager bemoans the fact that every year he must recruit staff from abroad because of the lack of US citizens skilled in math. He points to a shortage of qualified math teachers as national problem. Simons has founded a non-profit called Math for America “in an effort to train more qualified mathematicians” (Dobbs, 2004, A02) and is an accomplished mathematician himself. But he is identified only in his financial capacity, not by any experience he has as a mathematician or in education.

Elites as Experts

The speaker analysis provides insight into the profoundly elite nature of the PISA discourse. One might expect this, given that the perceptions of elites are particularly important in policy outcomes (Kingdon, 2011; Dahl, 1958). It also could reflect the fact that only about 5,000 students in the United States take PISA, suggesting that few

teachers, families and students are actually affected directly by PISA. Given the widespread agreement on PISA as a measure of national educational excellence, it is notable that only a handful of elites are given voice to comment on PISA. Government officials are privileged, as are the OECD and economists. Teachers, parents, students and school and local leaders are among the least likely to be quoted.

Favored individuals. Within the elite institutions favored in the discourse, certain speakers are quoted over and over again. In fact, the six most frequently quoted speakers provide more than one-quarter (26%) of the commentary. Figure 4.18 shows the most frequently quoted speakers, those that are quoted more than once throughout the discourse. The OECD, most usually in the form of Andreas Schleicher, appears in 10 separate articles; current US Secretary of Education, Arne Duncan, appears in 8 articles; and the next four speakers⁶¹ appear in three articles each. There are nine speakers who are quoted in two articles each, and the remaining sixty-four speakers appear in one article each.

Favored individuals by source. Each media outlet favors certain speakers over others. *The New York Times* and *The Washington Post* prefer sources from the US government (the Secretary of Education, or a department of Education official), while *USA Today* rarely quotes government officials. *USA Today* is the only source that

⁶¹ They are Chester Finn, Jr., President of the Thomas B. Fordham Institute; Jack Jennings, founder of the Center on Education Policy; Tom Loveless, Non-resident Senior Fellow at the Brown Center on Education Policy at Brookings; and Mark Schneider, visiting scholar at the American Enterprise Institute and Vice President at the American Institute of Research.

includes speech from parents or teachers. *The New York Times* quotes the OECD at a significantly higher rate ($p=.001$) than other speakers, while *The Washington Post* favors a variety of commentators from unions and other education NGOs. All media give voice to university affiliates fairly evenly.

Figure 4.18: Speakers by numbers of speech acts

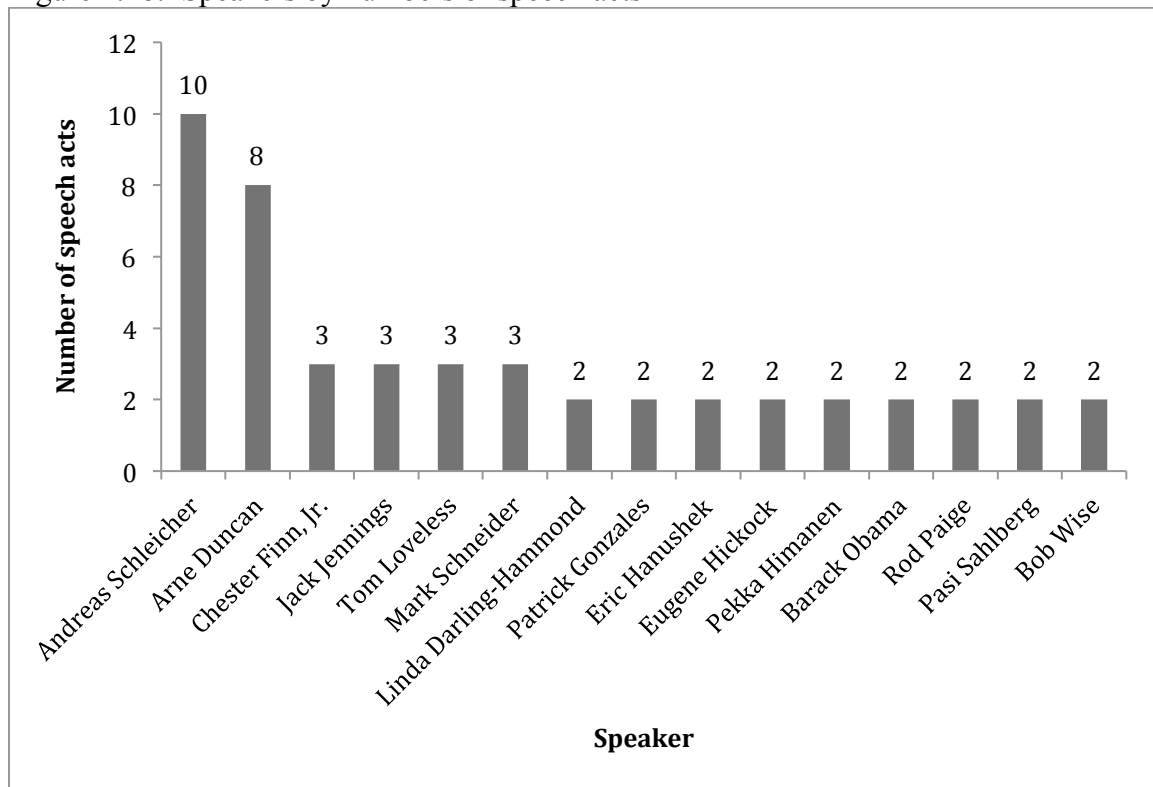
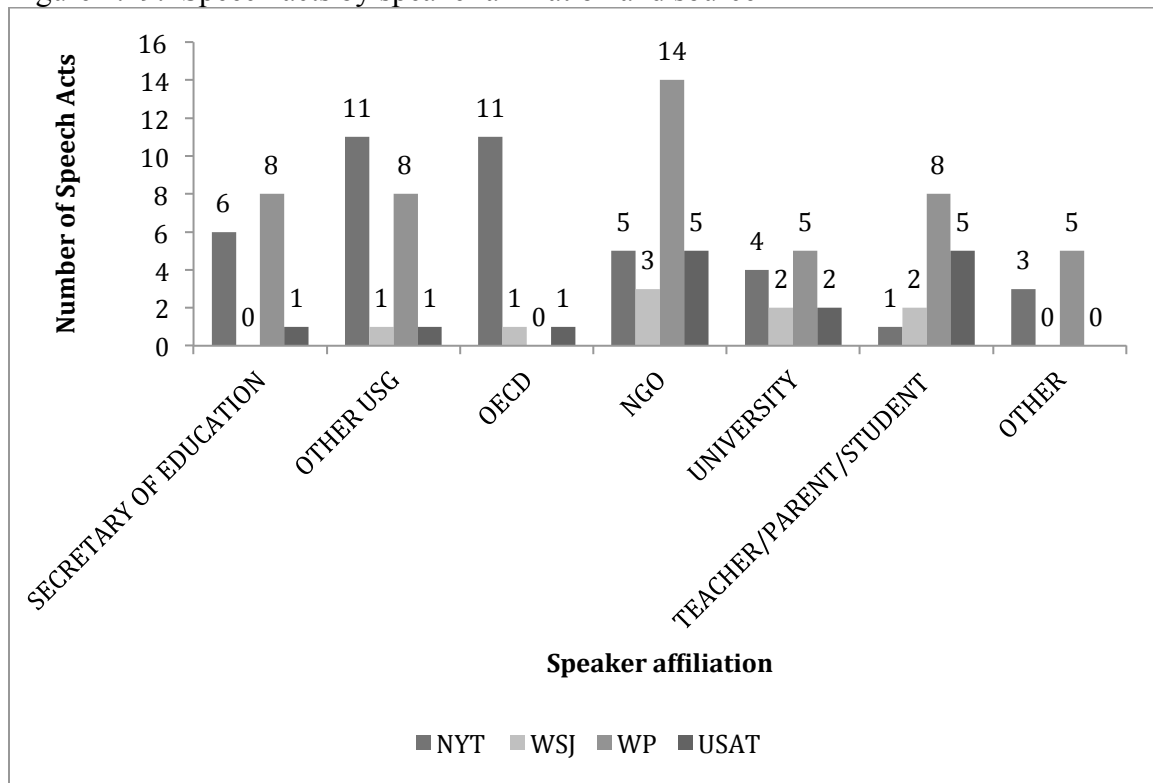


Figure 4.19 presents the relationships between sources and speakers, laying out which sources use which speakers in their coverage. Of the fifteen speakers that were quoted in more than one source, eight were quoted in two different sources. For instance, Andreas Schleicher was quoted 11 times in *The New York Times*, and once in *The Wall Street*

Journal. Arne Duncan was quoted six times in *The New York Times*, and eight times in *The Washington Post*, but never in *The Wall Street Journal* and only once in *USA Today*. The only speaker to be quoted in more than two media outlets was Chester Finn, Jr., whose “Sputnik” editorial was printed in its entirety in *The Wall Street Journal* and was also quoted in *The New York Times* and *The Washington Post*. Speakers quoted in *The Washington Post* are also often referenced in one or more other media outlets. This is largely due to the high overall share of articles about PISA published in the *Post*. *USA Today*, the least elite news outlet of those in the sample, has the fewest number of shared speakers, with only two men, Jack Jennings and Bob Wise, of the Center on Education Policy and Achieve, respectively, quoted in both *USA Today* and *The Washington Post*.

Figure 4.19: Speech acts by speaker affiliation and source



The OECD and the media

Although the discourse concerns domestic education policy and performance, the OECD is given outsized standing in the discourse. The coordinator of PISA at the OECD, Andreas Schleicher, is quoted in 10 articles, most frequently in *The New York Times*, commenting on all aspects of PISA and its results. He defends not only the technical validity and reliability of the test, but also offers general policy advice in his commentary. While Schleicher does not offer policy advice to the United States directly, he intimates areas where the US might reconsider its policies. “Teaching in the US is no longer a high-status occupation” so Schleicher encourages the US to recruit high-performing graduates and support them like other top-scoring countries do (Dillon, 2010, A22). In a 2012 op-ed article in *The Times*, columnist Thomas Friedman praises Schleicher and his team for delivering a “dose of reality” as they use PISA rank comparison as a policy strategy (Friedman, 2012, SR1).

It is not only the frequency with which the OECD/Schleicher is quoted but also the content of what the OECD says that is significant. While Schleicher steers away from prescribing specific policy options, he does use an education policy frame in his comments about one-third of the time, which is almost double what other speakers do. Schleicher repeatedly states that PISA should be a guidepost for national education systems around the world, and refers to PISA as a measure of educational excellence significantly more than other speakers do ($p < .05$). The frequency with which he is

quoted and the consistency of his message undoubtedly helped catapult PISAs status to a taken-for-granted measure of excellence.

Given the OECD's institutional commitment to economic development, it is not surprising that Schleicher frames his arguments for education in terms of economic returns. Quotes from the OECD/Schleicher use economic references significantly more than other speakers ($p < .005$), and mention returns to individuals twice as often as macroeconomic references. Schleicher is also one of the only speakers to capitalize on the competition inherent in the rankings, and he directly employs references to Finland and China as systems to watch. In the aforementioned column by Tom Friedman, for instance, Schleicher is quoted as saying,

“Imagine in a few years you could sign onto a web site and see this is how my school compares with a similar school anywhere in the world. And then you take this information to your local superintendent and ask, “Why are we not doing as well as schools in China and Finland?”

Whereas only 4% of speech acts mention Finland and 2% mention China generally, Schleicher uses reference societies in 40% of the articles in which he is quoted. The only other speakers to use reference societies are in direct response to the 2009 results, when several commentators including the Secretary of Education and the US President mention China by name, and its PISA success as a threat.

Clearly Schleicher, representing the OECD, is one the key voices in the discourse. A 2012 interview with him begins, “In the world of international education, what Andreas Schleicher thinks matters” (Guttenplan, 2012, TF6). Schleicher goes on to describe the efforts of PISA to build “a bulletproof instrument” that limits “political arbitrariness.” Schleicher is careful to tread fairly lightly when directly commenting on US education policy but consistently frames education in private, market-oriented terms and uses economic references significantly more than other speakers do. Schleicher consistently markets and drives the idea of PISA scores reflecting future macro level and individual labor market success.

Role of Opinion Writing

Farrell and Drezner (2008) report on the effects of blogs on politics, arguing the skewed nature of the blog readership, especially by elites and journalists from other media, gives undue influence and salience to a few highly trafficked and highly rated blogs. Such focal point blogs act “as a transmission belt between the blogosphere and politically powerful actors,” affecting political debate by creating a “menu of interpretive frames for the media to appropriate” (Farrell and Drezner, 2008, 22). In other words, an idea that becomes embedded in the blogosphere has the potential to focus subsequent mainstream news coverage of the same issues.

This seems to be the case with discourse around PISA in Phase II, when blogs and op-eds play an increasingly important role in the discourse. The fact that the blogosphere

responds forcefully to the 2009 results may have further encouraged the publication of articles and op-eds in the mainstream media. The most policy-oriented outlet, *The Washington Post*, posted a whopping 23 blog posts related to PISA, 85% of the total blog posts during that period. *The Post's* education blogger, Valerie Strauss, uses her blog to post not only her own views, but to re-post commentary gleaned from other online sources. *New York Times* op-ed columnist Thomas Friedman, who covers foreign affairs, devotes six columns to PISA related discussions during the same period.

Elites incorporate PISA into their own rhetorical repertoires, encouraging and perpetuating dominant understandings of PISA when they are quoted in the press. The heavy concentration of PISA discussion in the realm of opinion rather than news highlights the way in which education insiders appropriate PISA scores as evidence in their arguments. It also means that this coverage is not necessarily held to the same journalistic or fact-checking standards than a straight news article would be. The speed with which blog posts are published also allows for a virtually real-time discussion of the rankings and the score release each cycle. The concentration of PISA discourse in the realm of opinion rather than news suggests that education insiders are appropriating PISA results to support their own particular viewpoints, which I show in more detail below.

PISA as a Canvas for Education Elites

There is evidence that PISA is being used as reference in discourse by a variety of actors to try to gin up calls for reform. Elites use American performance on PISA as evidence

in the ongoing policy discussions about how US education should be improved. What is fascinating about the embrace of PISA by the education elite of the blogosphere is how commentators use PISA scores as evidence for any number of policy reforms. Following Smithers' (2004) and Waldo's (2012) metaphors of PISA data providing a blank canvas onto which actors can project anything they want to see, Figure 4.20 provides examples of ways different actors have appropriated PISA for a wide variety of claims.

The problem with this range of rhetorical claims is that it has meant that PISA has thus far received little traction as evidence for the need of specific policy reform. To date, PISA has meant something different for everyone, and there has been no clear agreement nor no new coalitions formed as a result of US PISA performance. While PISA is useful rhetorical evidence, it has been less powerful as a lever for policy reform.

Figure 4.20: Examples of PISA projections

PISA projection	Source
PISA scores “offer too limited evidence on which to base policy prescriptions...that is why more than 30 civil rights, education and disability groups...call for a shift away from the overwhelming reliance on standardized tests”	Op-ed in USAT, by staff at FairTest, a test reform NGO in Massachusetts
“Decentralization, competition and flexibility are on the curriculum suggested by Mr. Schleicher. We’d give those proposals an ‘A.’”	Editorial, WSJ
“As our test scores trail those of other industrialized nations, it is self-defeating to tolerate policies that impede assessment or inhibit innovation...we need better entries and exits from the system” (i.e., upgrade the teacher talent pool and decrease tenure)	Op-ed, Washington Post, by EVP for Public Affairs at Teach for America
“The United States’ scores on PISA don’t represent a crisis...but when it comes to the record of two-decades of test-based accountability reforms...[the scores show] its not working.”	Re-post on Strauss education blog, Washington Post, by anti-charter school academic
“The rankings from ...PISA are out, and our standing in the international comparisons has sparked debate on how to improve public education... among the hallmarks of high-performers such as Canada and Finland, are strong teachers unions”	USAT op-ed, by the president of the largest teacher’s union in the United States

Conclusion

Most of the PISA discourse takes place in the media. As in the academic and the think tank literature, most of the references to PISA in the discourse are simple and, as in the other literatures, the media discourse shows a lack of critical perspective. There is limited critique of PISA, either as an instrument or as a measure of educational success. Instead, there is broad agreement throughout the media that PISA is a bellwether of educational excellence especially in the context of the interconnected, global economy.

The media discourse focuses on improving PISA rank in the service of economic growth, rather than for the sake of improving teaching and learning. The media discourse is profoundly elite; the discourse is located in the elite rather than the popular press and most speech acts come from a handful of education policy elites. Almost 30% of quotes in the discourse originate from the same six speakers. There is very low public voice in the discourse. When parents, teachers, students or school leaders have a voice, it is when their writings are posted as guest blogs on *The Washington Post's* education website. If blogs are considered as a form of opinion writing, the data show that 64% of the PISA media discourse is published not as news, but as opinion.

During the decade of testing under review, the tenor, framing and voice of the media coverage shifts dramatically. The discourse is such that geo-economic and geopolitical relationships drive coverage. China's top scores on the 2009 PISA cycle recharge the discourse and push it towards a crisis narrative. As predicted by the externalization framework, the media articles reference foreign examples to create crisis, and to frame and strengthen arguments about the state of American education.

Theoretically, reference societies act as legitimating comparisons in contentious policy discussions. Actors use comparisons to foreign examples instrumentally to advocate for particular positions or policies, especially those that are contested domestically (Steiner-Khamsi, 2004). However, it is only at the rhetorical level that PISA has any bite; despite the narrative of decline that develops around PISA in the US media, PISA rank has only

been used at the level of policy talk, not policy action. This is markedly different from the reception of PISA scores in other country contexts. In order to understand how PISA reception varies across contexts, in the next section I look at the US reception of PISA comparatively, looking at the United States alongside the cases of Japan and Germany to elucidate the variables that constrain and facilitate PISA response.

CHAPTER FIVE:
PISA in comparative perspective:
Reception in Germany, Japan, and the United States

As I demonstrated in the previous chapter, the United States' reception of PISA is a subtle mix of embrace and indifference. American education elites embrace PISA in their discourse, and use PISA as a way to measure educational excellence. But the same elites are indifferent to actually using PISA as a policy learning or policy transfer device.

When comparisons to top-scorers are made, American analysts quickly conclude that in fact there is little to be learned from other countries, because other countries are too different from the US to really be models for it. There is no substantial lesson drawing for the United States from the experiences of PISA top scorers. I also find that references to PISA are not used to justify the contentious implementation of the Common Core State Standards. According to the externalization framework, this moment of contentious policy change would be precisely the time that references to PISA would serve to legitimate the contested policy action at home.

However, accounts of PISA reception in other countries demonstrate alternative trajectories. How has PISA been used as evidence for policy change in other contexts? In this chapter I begin to address this question by analyzing the US case in comparative perspective. In order to understand the factors and circumstances that effect PISA reception, I use secondary source materials to construct accounts that look at PISA reception across countries. I aim to provide a more comprehensive look at the US from a comparative perspective.

In this chapter, I analyze PISA reception in three country contexts: Germany, Japan and the United States. For the US case I draw on my original data and analysis from the previous chapter. Though a total of 70 countries participated in PISA during its first decade, I purposefully focus my comparison on these three countries for several reasons. Each country has participated in PISA in all rounds of testing. Each also represents a diversity of systems (centralized v. decentralized); educational cultures (Confucian v. western); and geographical locations (Asia, Europe and North America). In addition, the US has interesting historic ties with each: both countries have been viewed at varying times as models of education (Germany/Prussia in the 1800s, and Japan in the 1980s; see LeTendre and Baker, 1999) and economic rivals. Most importantly, both Germany and Japan have reacted to PISA, but the type of reaction, the timing of the reaction and the actors involved in each case have varied. By engaging in a systematic comparison of the three cases, I aim to make a preliminary step towards understanding when and how PISA affects national education politics.

After briefly discussing the comparative methodology and its limitations, I proceed by discussing the discursive impact of PISA in each country. I then explore possible reasons for this discursive impact, addressing four potential explanatory variables. I end with a short discussion and conclusion.

Comparative Case Study Methodology

The case study is a useful research strategy for identifying the processes through which conditions get translated into outcomes for future hypothesis testing (Van Evera, 1997). A comparative case design is appropriate in this instance in that it serves to help better-specify the mechanisms by which PISA may influence national education policy discourse in different country settings. Case study work can help to generate new variables and/or uncover mechanisms through which explanatory variables influence the outcome of interest (see Sambanis, 2004). Germany, Japan and the United States are offered as cases to begin to identify the circumstances under which ILSAs affect national education policy.

The comparative analysis is conducted using my original dataset of the US case and secondary sources describing PISA reception in Germany and Japan. I develop a bibliographic sample for Germany and Japan, and draw on extant English-language literature to analyze each country. Additional sources are drawn from the broader PISA literature, from which I conducted more bibliographic sampling to search for appropriate data. I also searched the online educational database ERIC for additional journal articles and references.

Limitations

As I discussed in Chapter Three, small-N comparative studies have their limitations, and one that relies on secondary analysis that much more so. The available data and the

number of cases are both constraints on the comparison. In particular, relying on English-language descriptions of PISA reception limited the available data. Not only were the available data constrained by language, but also those articles that did discuss PISA reception did not always have a research design similar to my approach in the US case. This led to missing data in some instances. Also, the role of the academy, think tanks and the press are also quite different between the three countries. In Germany, for instance, academics are considered public intellectuals and are often quoted in the press, while political parties tend to have a stronger role than think tanks in political discourse. In Japan, think tanks play a minimal role in the discourse compared to politicians. Fortunately, several of the same variables in the extant literature were the same as those used in the analysis of the American case, facilitating comparison across cases. While not ideal, conducting a secondary analysis allows me to overcome language limitations that might otherwise have made a between-country comparison prohibitive.

Capturing the Discursive Effects of PISA in Three Countries

To capture the discursive effects of the PISA discourse in Germany, Japan and the US, I look at how PISA is received in the academy, in think tanks, and in the media in Germany and Japan. In the German and Japanese cases, the data primarily covers media reception; there is little data on academic reception and almost none on think tank reception. For comparability purposes, then, I chose to limit my analysis to media reception across the three countries.

Using media reception as the outcome of interest, I categorize the PISA discourse in each country along three dimensions: (1) the quantity of media coverage; (2) the tone of media coverage and its use of reference societies; and (3) the effects of the media coverage on specific policy debates.

The quantity of media coverage. This category refers to the reported increase or decrease in mentions to PISA performance and ranking results in the national media. In the US case, this data is drawn from my original sample of four media sources. In Germany and Japan, coverage is extrapolated from broader accounts of PISA reception.

The tone of media coverage. I analyze tone according to the degree to which the narratives about PISA are accepting of the home country's PISA performance. I use a continuum ranging from glorification (the media glorifies the home country's performance); indifference (the media report on PISA in a limited and/or neutral way); and crisis/scandalization (the media characterize the home country's performance as scandalous and/or use PISA performance as indicative of an educational crisis in the home country), à la Steiner-Khamsi (2003). I also assess whether the PISA narrative includes references to top-scoring countries, and if so, which countries are referenced.

The effects of PISA on policy. This category refers to the discursive use of PISA scores as evidence in specific policy debates. Here, I look for the use of references to PISA scores by political actors as evidence to support particular education policies. PISA

performance becomes the part of the explanation as to why specific policies should be adopted, dropped, modified or continued.

Taken together, these dimensions of PISA reception begin to unpack the variation experienced across national contexts. Table 5.1 provides an overview of PISA reception across Germany, Japan and the US. I discuss each of the cases below.

Table 5.1: PISA reception in German, Japan and the United States, 2000-2009

	GERMANY	JAPAN	UNITED STATES
QUANTITY OF COVERAGE	Increased	Increased	Increased
TONE	Crisis and scandalization	Crisis and scandalization	Early indifference to contained crisis
REFERNECE SOCIETY/IES	Finland Sweden	Finland	Finland China
TIMING	PISA 2000	PISA 2003	PISA 2009
POLITICS AND POLICY	PISA scores linked to major reform Agreement across political parties	PISA scores linked to major reform Contention across parties	PISA scores not linked to reform Agreement across parties

Reception in Germany, Japan and the United States, 2000-2012

Germany. *PISA coverage.* By all accounts, the reaction to Germany's 2000 PISA results was a full-blown crisis. The resulting 'PISA shock' was felt across Germany, as the results were compared to Sputnik and the French Revolution (Ostermann and Hermann in Ertl, 621). By one count, regional newspapers in Germany published, on average one article per day for more than a year after PISA results were

released in 2001 (Tillman, in Ertl, 631). Another study showed that from the release of the 2000 results through 2008, the German daily newspaper *Süddeutsche Zeitung* published 253 articles about PISA (Martens and Niemann, 2010, 15).⁶²

Crisis and scandalization. The German reaction to PISA 2000 scores was crisis and a scandalization of German performance. Germany's scores were below the OECD average and Germany was ranked 25th out of 32 countries in reading and 21st in math and science. Germans were shocked at the socioeconomic achievement gap; poorer children performed significantly worse than better-off kids (Rubner, 2006, 255). According to Ringarp and Rothland (2010, 423), the disconnect between self-image and performance gave rise to a sweeping national discussion about how *das Land der Dichter und Denker* ("the land of poets and thinkers") could have gone so astray. Tröhler (2011) emphasizes this gap between self-understanding and results as a deep trigger behind the fierce German PISA reaction. He argues that the way PISA marginalizes knowledge (*Bildung*) in favor of competence is deeply unsettling to German understandings of education. Given Germany's cultural, ideological and historical background, Tröhler writes, "PISA's advance feels like an unfriendly takeover." (Tröhler, 2011, 250).

⁶² In light of the PISA shock German academics also reoriented their discourse towards a more empirical approach. Bohl (in Ertl, 2006, 627) reports that prior to the PISA uproar academic analysis had tended to adopt an approach where educational experiences were seen as personal and unique. PISA turned this approach on its head, so that everything in education became 'comparable.' A growing number of books and articles focused on the new educational changes. A search for "PISA study" in the German National Library catalogue returns more than 150 hits, and another search of the scientific database for pedagogy returns 167 hits for PISA 2000 and another 134 hits for the results of the PISA 2003 study (Ringarp and Rothland, 2010, 423). According to these accounts, the academic discourse grew enormously in tandem with the general public's focus on PISA results.

Reference societies. The most favored reference societies in the German discourse were Finland and Sweden, as they both represented models of equity and achievement. Geography made both countries ideal objects of affection, and the release of PISA 2000 scores spawned a significant eddo-tourism industry to both countries. The ease with which both countries could be visited only helped to support German reformers' fascination with them. Finland was characterized as another 'education utopia' (Takayama, Waldow, and Sung, 2013, 7). Finland was a favorite because of its top performance and the equity of its outcomes, as well as being the country with the smallest achievement gap between the highest and lowest performing students (Amermuller, 2004). Finland became a sensation, and followed the same ebb and flow of general PISA news coverage in the press (Takayama, Waldow, and Sung, 2013).

Ringarp and Rothland (2010) describe Germany's longstanding fascination with the other "Northern darling," the Swedish education system. Germany had long considered Sweden a "pedagogic wonder to the north" (Ringarp and Rothland, 2010, 424). This longstanding admiration was reactivated during the German PISA crisis and led to teachers, union members, researchers and politicians visiting Sweden on study tours. These visits produced numerous articles and reports, but no policy borrowing. Often Sweden was lumped together with Finland as part of the "Scandinavian success" in education, even though Finland is not a Scandinavian country.⁶³ The Germans saw

⁶³ Though geographic neighbors and a Swedish kingdom until 1809, Finland is not part of Scandinavia. Denmark, Norway and Sweden are considered Scandinavia, bound by history, geography and the same

Sweden's educational system as one where equal opportunity and high achievement were equally embraced and glorified Sweden as a model in public discourse. Ironically, the same educational policies that the Germans were glorifying at home were actually becoming highly contested in Sweden by the Swedes themselves, who felt that their education system needed reform. They began to turn towards a more centralized, assessment oriented system.

Takayama, et al., (2014) have paid attention to the ways in which "Asian" countries have been referenced in the German PISA discourse. They note that Asian countries are often lumped together in shorthand for a particular type of education that includes rote learning, excessive drilling, outside-of-school coaching, and emotional distress. Employing these stereotypical images sends the message that in Asia, good results come at a high cost. Takayama, et al. (2014) argue that one of the reason Japan did not become a reference society for Germany was that the methods by which Japanese students succeeded (cramming, feeling miserable) were deemed undesirable. Thus, while Asian countries are present as a reference point in the German discourse, it is the rejection of their approach to education that both the left and right agree on. According to Takayama, et al., (2014) Asian education, and especially Chinese education, is portrayed in Germany as an educational horror, and a kind of learning dystopia (p. 7). This kind of negative referencing turns out to be a powerful discursive move in the discourse, following Schreier and Martinez's (2004) argument that reference societies are used

language group. Finland is often included with those countries as part of the Nordic countries, as a geographic affiliation.

because they represent some aspect of educational society that actors wish to cleave to or distance themselves from. In the German case, Asian systems are referenced as the negative image of the reforms Germany should pursue.

Links to policy. In Germany, the PISA shock of 2001 was tied to far-reaching and concrete policy change. Critically, the political environment was primed for change; the Länder (state) and the federal government had been discussing school improvement reforms for some time. Germany participated in TIMSS in 1997 and was outperformed by 20 of the participating countries. According to Lehmann (2011), the reaction to the results was one of shock by both politicians and the public. The German conference of state education ministers (*Ständige Konferenz der Kultusminister der Länder*, or KMK) moved quickly to introduce reforms in the form of national standards and standardized testing. The so-called Konstanzer Beschluss (Constance Resolution) established that Germany should continue to participate in international comparative studies regularly, including PISA 2000. When the PISA 2000 results were released in 2001, and German scores were again mediocre, the shock only grew. Politicians became fully galvanized and resolved to make the early reforms happen, and the new standards and testing regimes were passed (Schmidt, et al., 2009, 12). On other issues such as quality assurance, for instance, the Länder had been looking at intensifying programs, both at the individual school and state administrative levels since the early 1990s. These initiatives moved forward rapidly after the release of PISA 2000 results (Döbert, 2007, 310). As one analyst notes, once PISA 2000 results were released “education policy making

without reference to internationally established norms was no longer a serious option” (Lehmann, 2011, 420-21).

Japan. PISA coverage. In Japan, the turn toward scandal and crisis occurred with the release of the PISA 2003 results at the end of 2004 and was intimately tied to ongoing political debates about the implementation of the *uteri* policy. Beginning in 1998, the MEXT had announced a revision of the national curriculum standards. These reforms included a decrease in instructional time, the streamlining of curricular content, the introduction of an integrated studies course that cut across traditional academic disciplines, and a five-day school week (Bjork and Tsuneyoshi, 2005); these were touted as *yutori* or relaxation of the high-pressured atmosphere that had come to characterize Japanese education. In 2002 the new education reforms were implemented, but became a political flashpoint when Japanese students did not perform as well as had been expected on the PISA 2003 cycle, despite the fact that it would have been highly unlikely that a policy implemented only one year prior to PISA would have impacted scores in the way in which it was portrayed in the press.

Takayama (2012) shows the media as playing a key role in framing and promoting a crisis discourse around Japanese PISA results. He analyzes three newspapers from diverse political views and finds that the Japanese media behaved similarly to the German and American media in terms of embracing a crisis discourse and favoring

speakers who were engaged in the politics of crisis.⁶⁴ Japan's scores dropped in all subject areas from the first round of testing in 2000, though Japan was still outperforming the OECD average. Regardless of the fact that its scores were not significantly worse, the media emphasized an educational crisis. He also argues that the Ministry of Culture, Sport and Education (MEXT) had no choice but to promote an anti-*yutori* policy once the public media narrative was out.

Crisis and scandalization. The opponents of reform asserted that the old policies were the very reason behind Japan's post-war rapid modernization, and that the reforms were threatening the academic standards that had been carefully developed over the years and had resulted in a high performing system (Bjork and Tsuneyoshi, 2005, 622). The drop in PISA scores in 2003 legitimized the need to revise policy again (Kameda, 2013). Takayama (2007) writes convincingly of the construction of the crisis of Japanese education with the appropriation of PISA and TIMSS scores. The media discourse presents the scores as a highly politicized debate between conservatives, who use PISA performance as evidence for a need to return to the stricter educational policies of the

⁶⁴ Takayama also notes the role of academics in Japan. He argues that scholars in Japan welcomed the PISA project because of the trove of good data it provided. Previously, as in most countries, there had not been such a detailed, freely accessible data on student achievement as well as student background and those in the academy began to use the new data to ask different questions. As in the United States, Japanese scholars were especially focused on equality, and used the data to create undertake research that focused on socioeconomic disparities and achievement. In so doing, they developed a narrative about the potentially negative consequences of the education reforms on equality. In Japan, these progressive scholars played a critical role in the politics of PISA, but also had very little influence in the mainstream discourse because their ideas never circulated outside of academic journals (Takayama, 2012, 159).

previous decade, and the progressives who argue that PISA performance has nothing to do with the new, more humane *yutori* policies.⁶⁵

Reference societies. Finland, as in Germany and the United States, is a preferred reference society in the construction of Japan's education crisis. In Japan, progressives glorify Finland as a symbol of opposition to both the conservative agenda and to what they themselves find wrong with Japanese education. Finland is portrayed as the model of a benevolent social-democratic state in opposition to the favored neoliberal position in Japan. The progressives emphasize that Finnish success comes from its commitments to a strong social welfare system, a high quality of life for all citizens, and social equity. Finland is able to become the "other" against which Japanese education policies are scandalized (Takayama, 2010, 18). Japanese neoliberals also embrace Finland. The neoliberals see in Finland another small state that has succeeded by relying on market principles, competition and choice (Takayama, 2012). Finland is a less attractive reference for Japanese conservatives: neoconservative reformers prefer to use the United States and Britain as their reference societies of choice to argue that Japan is wrongly moving away from the Anglo-American model.

⁶⁵ In a fascinating twist, Takayama (2008) shows that the media in Japan make extensive use of the crisis language from *A Nation at Risk*, a report that was written in large part as a response to Japan's perceived economic and educational dominance in the 1980s. The irony is extended as the United States has consistently scored significantly lower than Japan and the international average across all cycles of PISA.

Links to Policy. In Japan, PISA scores provide the political cover that Japanese politicians need to change direction in a contested policy debate. MEXT in particular leveraged Japanese performance on PISA to legitimize an about-face on *yutori* reform. By 2003, the Ministry was under pressure not only on *yutori* policy but also in moving towards decentralization and privatization in the whole education system. According to Takayama (2007), Ministry officials saw that they could use PISA scores as an external source of legitimacy. They acknowledged an education crisis and then re-oriented the highly contentious *yutori* debate. The strategy worked. Within a few months of the announced *yutori* reform in 2002, the new Minister of Education announced his intention to undo many of the key efforts of the reform package, using Japan's PISA performance as evidence for the need for change. Debate about the degree to which Japan should have a *yutori* strategy continued as a key issue in education policy throughout the 2000s (Gordon and LeTendre, 2010).

The United States. PISA coverage. As I have shown previously, the United States has a fairly indifferent reaction to the first three rounds of testing. Throughout the discourse, coverage of PISA does grow, but Phase I (from 2000-2009) sees relatively moderate coverage in both quantity and quality. The US performs below the OECD average on the early rounds of testing, but hysteria does not ensue. In 2003 and 2006, media coverage offers various caveats that US scores might be worrisome but still no alarm bells sound. Then the coverage explodes in after the release of 2009 results, when China is ranked first.

Contained crisis. It is not until 2010 – coinciding with China’s top scores on the 2009 cycle -- that PISA scores in the US are characterized as a critical measure of US educational success. Only then does the discourse turn scandalous, and PISA receives serious coverage in which alarm bells about America’s international standing are sounded. After flurry of crisis talk, PISA is remains a part of the American education discourse, but does not provoke any ongoing crisis, or point to any top-scoring countries as models for educational reform.

Reference societies. China is the important “other” in the American case. As I have shown, after the 2009 results are released, the entire discourse changes. Given the geopolitical relationship between the US and China, the media and education elites interpret China’s scores as a major educational and economic threat to the US. Even here, however, there are no actual calls for lesson drawing or policy borrowing in the discourse. Instead, references to China and Finland rise, and PISA is incorporated more widely as a reference throughout educational discourse. But the effects remain discursive and contained.

In Germany and Japan, Finland is a preferred reference society not only because of its high scores, but also as a symbol of equity and, in the case of Japanese neoliberals, market-oriented reform success. In the US, Finland is primarily referenced simply as a ‘top-scorer.’ When it is discussed more in depth, Finland is hailed as a model in terms of its teachers (their training, pay, and the respect they are given) and also in terms of its

commitment to equity. The most frequent reference to Finland in the US is as a model of equity, as it is in both Germany and Japan. Finland's ethnically homogenous population and strong welfare state support its educational success. Neither German, Japanese nor American actors suggest overhauling their national education policy in any serious ways to make their education systems more like Finland in these regards.

Links to Policy. In the United States, PISA scores have not been used as evidence to justify concrete policy action. I have mentioned the development and implementation of the Common Core State Standards as one contentious domestic policy action that could have been easily tied discursively to PISA. Though a cursory examination of background documents about the development of the CCSS often mentions the fact that domestic standards should be benchmarked to international standards, there were few links between PISA and the CCSS in the literature I reviewed. The American think tank literature shows that the PISA instrument was used in part as a model for the CCSS as they were being developed (Carmichael, et al., 2009), but the PISA discourse does not include references to the Common Core. Broadly speaking, PISA fits as part of the trend of data-driven accountability that has become increasingly prevalent in the US during the past decade. But in the texts I reviewed, PISA is not a strong part of the public discourse around policy reform.

What Explains Variation in PISA Reception?

From the brief descriptions of media reception in Germany, Japan and the United States, there are several areas of similarity and of difference on which to elaborate. The media in all three countries embrace PISA performance as newsworthy, and incorporate it into their coverage of national education issues. PISA performance is used to describe educational crises in each of the countries, though the timing and length of the crisis discourse varies. Each country uses references to top-scorers on PISA as a discursive strategy of comparison, though the reference societies themselves vary.

Drawing on my previous research as well as the existing literature, I explore five possible factors that might impact the reception of PISA across national contexts: (1) PISA scores themselves; (2) timing of the release of scores vis-à-vis the national policy process; (3) the actors involved in the PISA discourse; (4) federalism; (5) the way in which PISA is problematized in the discourse.

Scores. Table 5.2 provides mean PISA scores and rank for Germany, Japan and the United States compared with the OECD mean score on the first four rounds of PISA, between 2000-2009.

The table shows clearly that scores and rank themselves tell us little, if anything, about the timing, tenor, or reception of PISA. Of particular interest are those scores that are statistically significantly different than the average international score, as represented by

the OECD score. Germany scored significantly lower than the OECD in the first round of testing and is set off a major political crisis. In Japan, where scores across all subjects and all years were almost universally significantly *higher* than average, a crisis nonetheless surfaced around PISA 2003.

Japanese scores and rank did decline over time, with increases in both within school and between school variance (Park and Lee, 2013). However as Park and Lee show, the decline in performance scores was not uniform; the lower end (10th percentile) scores declined substantially, while the score for students in the 90th percentile remained virtually constant (Park and Lee, 2013, 134).⁶⁶ In the United States, where scores across most subjects and years were either average or significantly *lower* than average, there was no response until PISA 2009 – not because US scores had changed significantly, but because of who was first in the rankings.

Table 5.2 shows that actual rankings tell us little about the timing, tenor, or extent to which public discourse about PISA exists. Martens and Niemann (2010) raise this issue, noting that even countries that perform similarly vary in terms of public reaction. In the cases compared here, the United States has the most indifferent reaction, though it has the consistently lowest scores, while Japan, which performed consistently significantly better than the average country, has an educational crisis that results in policy change. The numbers themselves are ambiguous until they are constructed as symbolic of educational

⁶⁶ Park (in Park and Lee, 2013) observes similar trends in a separate study of math performance.

decline. Though the worldwide embrace of PISA supports the hegemony of empiricism, the actual scores themselves need to be assigned meaning before they become politically useful.

Table 5.2: PISA Scores for Germany, Japan, the US and the OECD, 2000-2009

Subject	Country	2000	2003	2006	2009
Mathematics¹	OECD	493	500	494	496
	Germany	490	503	504	513
	Japan	557	534	523	529
	United States	493	483**	474**	487**
Reading	OECD	494	494	489	493
	Germany	484**	491	495	497
	Japan	522	498	498	520
	United States	504	495	---	500
Science	OECD	494	499	498	501
	Germany	487**	502	516	520
	Japan	550	548	531	539
	United States	499	491**	489**	502
<p>** Score is statistically significantly lower than the OECD score, which represents the international average.</p> <p>¹Significance tests were not available for 2000 math scores.</p>					

Timing. The variation in cases suggests that timing in the policy cycle is an important variable in whether and how PISA performance data are used in local contexts. Returning to Kingdon's (2011) model, which portrays problems, policies and politics as largely independent, signaling events are one way in which the three streams can crystalize and come together. In the cases of Germany and Japan, PISA was used as

symbol of crisis for reforms that were already fomenting. This seems key to the fact that PISA impacted policy discourse. The release of PISA scores acts as the signaling event that aligns timing, public support, political agreement, and resources. But PISA's impact occurs when there has been ongoing policy deliberation, akin to Kingdon's 'softening up' period (2011, 127-130). The fact that the release of PISA scores occurred as ongoing reform debates were taking place in Germany and Japan, and where reform ideas were already keyed up, allowed PISA be easily incorporated as evidence into discourse.

In Germany, talk about reforms started in 1997, after the TIMSS results were released. TIMSS performance triggered German policy makers for look for solutions to what they saw as disappointing German TIMSS results. That concern began a movement toward standards and standardized testing to evaluate the Germany system. According to Schmidt, Hoang and Shakrani (2009), these reforms were at the time rather contentious, with the states offering strong opposition. Things shifted, however, after the PISA 2000 results were released. The similarly disappointing results on PISA were just the fodder that the reformers needed to argue for rapid implementation. With the discourse focused on the PISA results, German officials became fully galvanized to make the reforms happen (Schmidt, et al., 2009, 12).

Similarly in Japan, conservative politicians had begun trying to reverse the implementation of the *yutori* reforms as soon as they were put in place in the early 2000s. In Japan, the drop in performance scores from 2000 to 2003 provided excellent

ammunition for the conservatives' arguments. The tensions between a more relaxed approach and a return to traditional curriculum has continued throughout the 2000s, with politicians continually drawing on PISA performance as part of the Japanese discourse of reform (DeCoker, 2013; Gordon, et al., 2010).

In the US, the “crisis” of the PISA 2009 scores did not aligned optimally to create a discursive link. PISA 2009 scores were released in December 2010, and press coverage began immediately after that, with most of the crisis talk occurring beginning in December 2010 into early 2011. Though the CCSS were adopted by states in 2010, states had to commit to adoption by August 2010 if they were to be eligible for Race to the Top funding. Thus the PISA 2009 performance was in the headlines only after states had made their CCSS participation decisions.

Actors. The comparative case analysis suggests at least two potential ways in which actors are important to PISA reception. The first, as suggested by the literature, is that the prominence of the actors involved in the discourse affects the degree to which PISA is incorporated into policy discourse (Simmons and Kelley, 2015). In Germany and Japan, legislators and elected officials were deeply involved in the PISA discourse that helped to put PISA discourse in the center of education coverage in the media. In Japan, the Minister of Education and conservative legislators used PISA data to argue that the *yutori* policy was weakening Japan. In the United States, the discourse relies more heavily on OECD officials, and elite analysts from think tanks and other non-

governmental organizations. This may suggest that the involvement of elected officials and/or legislators is harbinger of PISA being used for specific policy purposes. As the general PISA discourse draws in policy makers, it is likely that those policy makers then turn to PISA to support specific policy reforms.

Secondly, PISA data may be used to bring together new coalitions of education actors who had not previously been aligned around a certain issue. Actors who may not have been previously associated with an educational reform use PISA as the critical data that allows them to coalesce around the new issue. Many actors responded publicly to the German PISA shock and a new educational discourse began across a huge range of educational issues. German politicians, of course, were in the fray, but trade unions, parents, and academics, amongst others, also responded. Ertl (2006) describes a situation where the employers' union and the trade unions, previously diametrically opposed on almost all policy issues, issued a joint statement in support of the educational reforms. A similar situation occurred between the federal and Länder (state) system, which had previously divergent agendas, but found themselves in agreement on the "frantic" reform agenda (Pongratz, 2006, 472).

Perhaps surprisingly, the PISA crisis in Japan helped to form a political alliance between neoconservatives and neoliberals, who held diametrically opposed viewpoints on the reforms. While the neoconservatives felt that the *yutori* approach was degrading patriotism and traditional values, the neoliberals had encouraged an emphasis on

individualism and state withdrawal. The crisis discourse helps to form an alliances between two factions of conservatives, who unite around the need for reform, even if they do not necessarily agree on the form the correct policies should take (Takayama, 2008, 2010).

In the United States, legislators and elected officials have contributed sparingly to the PISA discourse. Though there has been some commentary from governmental officials in the discourse, those elites are largely appointed bureaucrats rather than elected officials. The Secretary of Education is part of the contained crisis discourse, but does not tie American PISA performance to any specific calls for policy reform. The educational elites that are part of the American discourse instead project a variety of meanings onto the blank slate of PISA. But no arguments for reform have coalesced around PISA data or PISA performance scores.

Which American actors are in the discourse also may explain the lack of connection between PISA and the Common Core. The CCSS were developed by a private consortium of experts funded by private money; the US federal government played little *formal* role in the creation or marketing of the CCSS, though making RttT funding conditional on the adoption of standards gave the federal government serious leverage in the process. Since states adopt the CCSS individually, it is presumably state actors that will have the most to say about Common Core adoption and implementation. States,

however, do not as a rule participate in PISA.⁶⁷ Therefore at the state level, there is no immediate reason that state actors would either connect or act on PISA data and state policy.

Federalism. Federalism has previously been shown to affect the education policy-making process (e.g., Schmidt, et al., 2009). Germany and the United States have federal education systems where the locus of power resides in the states, though both have seen an increased role for national education authorities during the past decade. In each country, the locus of control seems to have minimal effect on the degree to which PISA was incorporated into the discourse or led to concrete policy change. In Germany, the Länder ministers had come together to work on a reform package in the late 1990s that was well underway by the time the PISA shock of 2001 hit. Low PISA scores helped to enact the new policies already pending. This was notable not only for the speed with which the political agreement took place but also because many saw it as an infringement on Länder autonomy, a prized characteristic of the German federal system. The Länder reformers turned away from the traditional German understanding of education though, and increasingly embraced empirical measurement and evaluation as a cornerstone of education policy (Döbert, 2007), for which PISA was the perfect vehicle.

⁶⁷ Interestingly, the OECD, along with an American NGO called America Achieves, has developed an OECD Test For Schools, which bypasses federal and state involvement in the assessment administration. Based on PISA, the test “enables individual high schools to benchmark their students’ performance against that of their international peers.” (see <http://www.americaachieves.org/oecd>)

In Japan, the PISA crisis was effective in bringing about more policy change not because it involved the Ministry of Education, but because it was used in a partisan debate that pitted the future against the past. Conservative legislators turned PISA scores into symbols of decline. Existing accounts of the policy change suggest that the fact that the system was centralized had less to do with the change in *yutori* policy than did the successful use of crisis discourse by conservative politicians (DeCoker, 2013, 7). Semantics and politics, rather than structural characteristics, put pressure on the policy debate.

In the United States, as connected to the discussion of actors in the previous section, the federal system may play a role in terms of how limited the effect the PISA discourse have been on national education politics. Despite the growing involvement of the federal government, policy implementation of initiatives like the Common Core occurs at the state level. PISA may feel too far removed from state policy makers for PISA scores to have any bite in more local policy contexts.

Problem definition. Scholars of policy have examined the ways in which a condition comes to be defined as a problem (Gusfield, 1981; Stone, 1989; Campbell, 2002; Mehta, 2011) and have highlighted the fact that narratives do not just exist as fact but are constructed. In Japan and Germany, powerful political actors use PISA strategically to create a narrative where PISA performance is a problem that must be fixed. In Germany, PISA 2000 scores, like TIMSS in 1997 before it, were specifically

characterized as emblematic of the decline of German education; national standards and other reforms already in the making stood as the perfect remedy. Policy makers and political actors from all sides agreed on this problem definition and came together behind the reforms.

In Japan, conservatives created a brilliant narrative around PISA performance, and used PISA scores as the evidence they needed to build political support for rolling back the year-old *yutori* policies. While conservatives and progressives may have disagreed on its meaning, the conservatives were successful in problematizing PISA performance and pointing the finger at the progressive policies as responsible. What was needed, as declining PISA scores demonstrated, was a return to the past.

In the US, PISA has not successfully been problematized in the same way. The American discourse does not provide a clear problem definition, nor are there strong claims about responsibility and blame. As I showed in the previous chapter with examples of projections (see Figure 4.20), there is no clearly defined problem that PISA is able to solve for American policy makers. PISA is thus far providing neither educational problems nor solutions. PISA's utility is also constrained by the fact that thus far there has been more agreement than disagreement about the meaning of PISA scores across partisan lines, though it is unclear if this agreement has led to any new reform coalitions as a result.

Conclusion

This comparative analysis illustrates some initial ideas about the general processes that could be at work in PISA reception across countries and suggests avenues for future research. My analysis emphasizes several points. Clearly, PISA scores explain little about PISA reception. The discourse is about politics, not empirical data. Neither scores that were significantly lower nor those that were significantly higher than the OECD average scores had predictable or similar reception across country contexts. In Germany, Japan and the United States, performance on PISA is received with dismay, though at varying times and to varying degrees. Despite this variation, actors in each of the three countries bestow symbolic authority on PISA so that PISA scores arbitrate the status of education. The scores themselves, especially whether or not the between-country gaps matter in statistical terms, barely register in the discourse. Country reactions are centered on ‘poor performance.’ But the crisis discourse is far more dependent on the constellations of timing, actors and narrative that interact to produce the environment in which PISA scores are evaluated and interpreted.

In Germany, the ‘PISA shock’ of 2000 catalyzed reformers of all political stripes to quickly address Germany’s wounded pride by pushing through a series of reforms that had been percolating for years. In Japan, where there was never any significant difference between Japan’s scores and the international average, Japanese politicians nonetheless used PISA to create both an educational crisis and a new round of reform. The US, which for the first three rounds of PISA was fairly indifferent to its own

performance, went into crisis mode after the 2009 results were released. However in the US the crisis remained at the discursive level, and has not resulted in any new policy change.

Who is involved in the discourse is also important. From the situations described in these case studies, PISA discourse is undoubtedly the purview of educational elites in all participating countries. But when legislators and elected officials take up the PISA narrative, it is perhaps more likely to be used as evidence in specific reform efforts. Policy discourse aside, it is clear that the PISA narrative in general is powerful across different contexts because it is numerical and data driven, which comports with common understandings of what scientific data looks like; it is versatile, and can be used as evidence of many educational problems and solutions; and is attractive to the media, because competition and rank makes for good news. This ensures that, at the very least, PISA will be an object of media interest in many, if not all, participating country contexts.

The comparative study also suggests that the way in which the PISA narrative is constructed matters. The use of reference societies is common to each of the countries but comparisons to top-scorers are made for contextual reasons. For instance, despite the fact that Finland is ‘objectively’ first in many of the league tables, not all actors find referencing Finland the best strategy in their domestic context. Each reference society resonates differently depending on the context in which it is used. Though it used

Finland as well, Germany turned to strongly to Sweden, an historic educational role model whose system was the opposite of Germany's – decentralized, autonomous, comprehensive, and inclusive. Japan's references were bifurcated, depending on the political orientation of the referee – neoconservatives turned to the US and Britain as models of tradition, while neoliberals turned to the elements of the Finnish experience as a model. Each country's use of reference societies is mediated by its goals for interpretation, its cultural traditions, political forces and dominant ideologies (Schriewer and Martinez in Steiner-Khamisi, 2004, 50).

The case studies also show that there seems to be no lesson drawing or policy borrowing in the literal sense. In Germany, policies that support a new regime of standards and accountability are implemented, but Japan turns to the US example as a source of external authority to bolster its own policy stance, borrowing the old crisis discourse of US to create its own education crisis at home. Takayama (2007) argues that regardless of whether these policies are actually implemented and/or working in the United States, the reformers appropriate the crisis language to support the construction of an education disaster in Japan. The extant literature suggests that the policy change that occurred in the aftermath of the PISA crises in Germany and Japan was not a result of educational policy transfer. That is, no actual policy borrowing from Finland or other top-scoring countries took place. Instead, the use of reference societies in Germany and Japan occurred as a discursive mechanism. To date, the use of reference societies remains purely symbolic. Despite the educational tourism to Finland, there are no accounts of

policy change taking place in any of the three countries because of an educational policy transfer. While successful countries are discussed as models because of particular approaches they use, none of the countries borrows policy wholesale. Insofar as there is any convergence, it is the *idea* of PISA success that is the common phenomenon in the three countries.

While these findings are necessarily preliminary and constrained by data availability and language limitations, they nonetheless suggest avenues for future research. Further analysis could systematically unpack the interplay of timing, actors, politics and policy response in additional participating countries. Here I have emphasized the constructed nature of PISA reception and meaning. It would be useful to look more thoroughly at how PISA is received in multiple contexts, since is meaning-making that is constitutive of the importance of PISA scores, not the other way around.

CHAPTER SIX: THE POLITICS OF PISA DISCOURSE IN THE UNITED STATES, 2000-2012

The empirical findings presented in the previous chapters elucidate the ways in which the PISA discourse has been constructed in the United States from 2000 to 2012. The data demonstrate that during this period there was a rapid, measurable incorporation of PISA into American educational discourse, primarily through simple, uncritical references in the media after 2009. The data show varied discursive representations of PISA, and several ways in which PISA scores are used as evidence. They also show how PISA is used to problematize multiple aspects of the American education system according to the interests and incentives of different actors and institutions. During the first decade of PISA, within all three of the literatures in the sample, PISA goes from being unknown and essentially ignored in 2000 to a taken-for-granted measure of educational excellence in 2012. Discourse in the United States was fairly minimal during the first three cycles of testing, but increased dramatically after the 2009 cycle results were released and China was ranked first in the PISA league tables.

Discussion of Findings

The US discourse pays scant in-depth attention to PISA. The vast majority of references to PISA in the dataset are simple, meaning they are the focus of less than 20 percent of the text. In many of the texts, references to PISA and/or reference societies are not more than a sentence or two. Sometimes the reference is no more than a phrase, a variation on “PISA, an international assessment administered by the OECD...” or “Finland, a

consistently top scoring country on PISA...” There is practically no discussion of what rankings mean empirically, or which ranking comparisons are significant in statistical terms. Nonetheless, the persistence of these simple references over time, coupled with a handful of high profile articles invoking ‘crisis’ around the 2009 results are enough to construct the idea of PISA performance as a rhetorical trope for education elites.

Importance of Timing

Examining the discourse over time shows how ideas about PISA are developed, both incrementally and instrumentally. In sampling literature from the first decade of PISA, I am able to show not only the growth of the discourse but the changes in the way PISA is constructed as evidence. Initially, the PISA narrative supports the longstanding view of mediocrity that has defined American educational discourse since the release of *A Nation at Risk*. Using the ‘hard’ numbers of PISA rank, elites continue this narrative with each PISA cycle, regardless of America’s empirical standing. More dramatically, elites use PISA data instrumentally, as a response to geopolitics. The claim that the US is in crisis vis-à-vis the Chinese marks a turning point by creating limited hysteria about US performance and cementing into the discourse the idea that PISA measures educational success. It also shows that reception is not static, in response to a particular PISA score or cycle of assessment. Reception varies over time. PISA data gains stronger purchase when it is politically useful.

Cleavages between literatures

PISA representation and use is not consistent across or between literatures. Academic literature is disconnected from think tank and media literature, while the think tank and media literature exhibit ties that shape the discourse in critical ways.

Academic literature. The American academy, as represented by AERA journals, remains isolated from the larger public discourse. None of the think tank or media articles cites academic literature or adopts ideas generated in the academic research. The academic work remains insulated, and has little exposure or authority in the larger public discourse driven by the press.

It is not wholly surprising that academics look to PISA primarily as a source of data. It is the job of the academic researcher to methodically work with data and unpack its meaning, and the incentives of the academy reward thorough and thoughtful review of data for publication. There has been a longstanding tension between the idea of scholars as producers of scientific, objective data in the pursuit of knowledge, and the idea of scholars putting that knowledge to work in the service of solving public problems.

Even with this tension, the academy is surprisingly accepting of PISA and shows little analytical work on the normative or cognitive effects of the PISA discourse. The comparative literature takes a slightly more critical view but, like the American research, remains isolated from the think tank and media literatures, as its cases are focused on

Asia and Europe with little scholarship focusing on the experiences of the US or other regions of the world. There is no crossover of ideas from the academy to American think tank or media literature.

Recent research suggests that scholarly engagement with public issues may strengthen scholarly credibility (O'Brien and Pizmony-Levy, 2015, in press). There have been calls for academics to try to have more of an impact in public discourse. Takayama (2012, 162) implores scholars to “circulate their ideas outside of conventional academic publication venues to intervene more effectively in the cultural politics” that can characterize PISA reception in national venues. *New York Times* columnist Nicholas Kristof echoes this view on a recent blog post that lambastes the academy for fostering “a culture that glorifies arcane unintelligibility while disdaining impact and audience” (Kristof, 2014).⁶⁸ Still other academics have suggested considering ‘altmetrics’ (Fischman and Tefera, 2014) to broaden the measures of scholarly impact on which tenure decisions are traditionally made. As far as the PISA discourse is concerned, however, the data indicate that academic impact in the PISA discourse has thus far been limited.

Think tank literature. Think tanks punch above their weight in the PISA discourse. Think tank texts are regularly quoted in the media, where the bulk of data is

⁶⁸ Kristof’s blog post ruffled more than a few academic feathers. For one rebuttal from political scientist Daniel Drezner see http://www.politico.com/magazine/story/2014/02/nick-kristof-academics-rebuttal103786_Page3.html#.VCgSakvMeds

situated. The analysis shows both Fordham and Brookings to be highly regarded sources that provide the media with commentary and data on PISA. Think tanks work to get their research noticed and leverage both institutional reputation and personal networks to do so (Rich and Weaver, 2000). Both Fordham and Brookings do remarkably well on this front. Both think tanks' research is picked up by the mainstream media and provides important framing for PISA coverage.

In particular, Chester Finn, Jr.'s opinion piece on the meaning of China's 2009 triumph almost singlehandedly spurs the crisis discourse that played out in the media for several weeks afterwards. Finn's provocative language was published initially on a Fordham blog, but also as an op-ed in *The Wall Street Journal*. *The New York Times* and *The Washington Post* quoted Finn as well. Finn's role as a leading voice in the PISA discourse comes with some amount of irony. On the one hand, Finn co-wrote the Introduction to the scathing Fordham report that compared Common Core standards with other assessments often used as benchmarks (e.g., NAEP, TIMSS and PISA). About PISA, he wrote: "PISA strikes out. Neither in reading (literacy) nor in math does its content deserve better than a grade of "D." This is not a promising benchmark for American K-12 education in these subjects" (Finn, C.E. Jr. and Winkler, 2009, 6).

But only a couple of years later, Finn sounds the alarm on American education by referring to PISA results. His blog post and op-ed are a lynchpin in the discourse, with even President Obama picking up Finn's Sputnik analogy in comments made at a

political event a few days after the 2009 results were released.⁶⁹ This speaks not only to the use of PISA as evidence in political rhetoric, but also to the importance of actor status in discourse. As a leading conservative education pundit, Finn is in a position to have his views disseminated widely, even if they contradict themselves.

Tom Loveless' work for Brookings also receives cross-platform coverage, cited by Fordham and the *Washington Post*, on both print and blog platforms. It is evident that both Fordham and Brookings are generating new ideas, using networks, diffusing those ideas and particularly targeting other elites. These relationships are typical of successful political think tanks (Gellner, 1995), which develop a deliberate closeness to the policy process as part of their missions. Brookings' and Fordham's abilities to leverage their ideas into relationships differentiate them from other types of knowledge producers. Fordham in particular is highly successful in disseminating its views in this discourse. Fordham analysts publish across Fordham platforms (print, online reports, blog posts) to create an institutional narrative around international assessment. The mainstream press use Fordham's research but not Brookings'. This supports previous research that conservative public policy think tanks tend to be more effective in terms of getting their ideas to mainstream audiences than liberal organizations (Rich, 2011, 191).

The orientation and language of the think tanks are different than the academy, and this helps think tanks achieve their goals. Where the academic literature takes an analytical

⁶⁹When the results from PISA 2009 were released, the US President issued a statement saying "our generation's Sputnik moment is back" (Obama, 2010).

view, the think tank literature focuses on a normative approach – what ‘ought’ to be with regard to PISA ranking. The think tanks are more likely to approach PISA holistically and do not always use raw data to answer specific research questions. They also use colorful rhetoric to take sides and assign responsibility and blame, giving the think tank literature a different discursive power. Without the constraints of neutrality and peer-review, and with the ability to leverage personal and professional policy networks for access and exposure for their work, Brookings and Fordham demonstrate an ability to get their ideas into the mainstream media.

Media. Media produce the vast majority of the PISA discourse and offer little critique of PISA, either as an assessment instrument or as an accurate measure of educational success. The media paid minimal attention to American performance in the first three cycles of PISA, but the narrative and tone of the coverage changed immediately in 2010, from that of a cautionary tale to hysteria over China’s ascent. The media appropriate crisis rhetoric from the think tanks that has a huge impact on both the quantity and tone of the discourse. Coverage explodes after the release of the PISA 2009 scores and continues to grow even after the crisis talk dies down.

The tone of the media coverage changes from news-oriented, descriptive coverage in the first phase of the discourse (2000-2009) to alarmist, after the 2009 release of scores. Though the crisis subsides rather quickly, the overall tone of the discourse changes.

PISA is incorporated as an important measure of educational, and economic, success.

PISA is referenced widely by elites in op-eds and blog posts as well.

Role of blogs and opinion writing. The comparative advantages of blogs are the speed with which their content is generated and the low-costs associated with their publication. Blogs can respond to political events in virtually real-time, meaning the discourse can be quickly disrupted with a new idea – as demonstrated by the introduction of crisis talk about the 2009 PISA results. This makes blogs a favored source of elite information, and a favored venue for discussion of ideas around PISA.

Farrell and Drezner (2008) point out the symbiotic relationship between journalists and political blogs, where personal networks and publication incentives create important channels for the exchange of ideas and information. The PISA discourse shows a movement and appropriation of ideas and frames between the think tank literature and the media, while the academic literature remains isolated. The use of blogs by academics is generally more suspect and more limited. Academics, who operate on different incentives and within different professional networks, tend not to leverage their academic research through blogs.

It is noteworthy too that blogs provide the major venue in which the ‘public’ (teachers, school leaders and parents) has a voice. *The Washington Post’s* education blog reposts blog entries from the sites of individual bloggers, many of whom are individual education

activists directly tied to schools. These posts provide the only media critique of PISA, though these voices remain marginal in the overall discourse.

Discourse is elite. The PISA discourse is an elite, not a popular, one. While this conclusion may largely be an artifact of the research design, i.e., that elite news sources will necessarily give voice to elite actors, the cleavage between elites and the public in the discourse is still dramatic. The cleavage between elites and the general public is visible at two levels, at the level of news source, as well as within articles by speaker. The elite press (*The New York Times*, *The Wall Street Journal* and *The Washington Post*) engages in the PISA discourse almost 12 times more often than the popular press does, as represented by *USA Today*. The coverage is not only more concentrated in the elite press, but the elite media cover PISA more consistently. My analysis shows that over time, articles are published not only as news events, around the time of the release of scores, but throughout the year, suggesting that PISA moves beyond being a news spectacle to having a taken-for-granted status, where references to PISA are made regularly in the context of general educational discourse.

At the speaker level, the discourse is guided by a mere handful of elites: the data show that six speakers provide almost 30 percent of the commentary on PISA in the US. Andreas Schleicher, the head of PISA at the OECD is the most often quoted expert throughout the US discourse. The majority of speakers are affiliated with the US Department of Education. Individuals from more than 40 non-governmental

organizations are also quoted, but the dispersion of quotes across many speakers and affiliations means that none of the NGO voices are especially strong. Speakers affiliated with universities are 11 percent of the speaker sample and most of them are men (even though seventy-six percent of American teachers are women); most of them are economists; and all are from elite institutions. Almost half (48%) of the 112 quotes in the sample come from the same 15 speakers, none of whom is a teacher, parent or student. Those closest to the schoolhouse have the least say in the discourse.

It is ironic that there is such low public voice in the discourse. Assessment is increasingly seen as *the* form of accountability in public education, yet the PISA discourse on education has no “public” in it. While PISA is not a major part of the US accountability regime, the power of assessment is extended when there is no public voice to offer a different perspective. The extent to which the public is closed out raises questions about to the degree to which journalists, editors and pundits are able and inclined to include a non-elite narrative as part of the discourse.⁷⁰ Given the limited space, time and money used to cover education issues in the media, the issue of whether and how the media should represent the ‘public’ remains an open one.

⁷⁰ This issue is not particular to my analysis. A recent list of “Top Twitter Feeds in Education Policy” published by Education Next caused consternation when its top ten list included only well-known and well-connected education personalities. The pushback led Education Next to revise its list to be more inclusive of those individual bloggers/tweeters it had overlooked. See <http://educationnext.org/the-top-twitter-feeds-in-education-policy-crowdsourced-edition/>

Discourse is Driven by Politics, not Evidence

The PISA discourse emerges around political, not empirical, concerns. PISA scores and PISA rank have little explanatory power in terms of whether they will result in domestic political backlash or policy change. The comparison of scores between Germany, Japan, the US and the OECD (Table 5.1 in Chapter Five) shows that Japan scored significantly *higher* than the OECD average in all subjects for all years and yet there was enormous political backlash to its PISA scores that led to a retrenchment of education reform. Germany also experienced ‘PISA shock’ upon the release of its early PISA scores, leading to the implementation of major new education policies.

The US has consistently performed significantly worse than the OECD average but there was no significant response until 2009, when China participated in PISA for the first time and was the top scoring country. American policy makers have long considered China a key foreign policy relationship, with China alternatively portrayed as both an ally and a threat in the media. Additionally, the two countries have a massive economic relationship – their trade and investment flows in 2012 reached a record half trillion dollars – and China continues to record near double digit annual economic growth rates (Ikenson, 2013). Against this deep relationship and the backdrop of the economic crisis of 2008, China’s top PISA performance brings newfound salience to PISA rank. US scores have not varied much over time, yet the response to the scores has changed dramatically. It is not the scores that vary, but the countries on the ranking lists. This

fickle usage suggests the political nature of the evidence and the importance of context in evidentiary use.

Use of References to Top Scorers

My analysis shows while Finland and other top-scorers are referenced in the first phase of the discourse (2000-2009), it is not until China is introduced that there is urgency in the discourse. The combination of pre-existing political and economic relations perhaps influenced by the serious economic downturn in the US conspires to make China's PISA scores a powerful rhetorical trope. As Germany and Japan were threats in earlier times, China is now the most vivid symbol of competition with the US vis-à-vis educational and economic success.

The China crisis discourse unexpectedly corresponds with a reactivation of references to Finland. Finland does not occupy anywhere near the same geopolitical or economic status as China within the United States but it is utilized as a counterpoint reference to China in the discourse. There is none of the hysteria that accompanies the Chinese discourse. Instead, Finland is most often described in the context of equity. The references to Finland are less frequent, and also simpler; usually they are references to Finland's status as a top scorer. Finland is no geopolitical or economic threat and therefore its references are fewer and pack less rhetorical punch. China is critical to the use of Finnish references, but the reverse is not true.

The reasons given for Finnish success are quite opposite from the depiction of Chinese success. Finland is portrayed as state where the strong social welfare system works to provide equitable access and results to all students. The coverage emphasizes characteristics like short school days, patient teachers, and lots of recess. On the other hand China, it is argued, encourages its students to perform well on international assessments as a matter of national pride. Students succeed through rote learning and cramming and the educational environment in cities like Shanghai is highly pressurized.

Additional Puzzles

My analysis raises many more questions than it answers. Most obviously, why hasn't PISA been used as evidence of 'best practice' in education? Top scoring countries seem to do something right educationally; perhaps their policies have something to offer. Yet there are almost no calls for lesson drawing or policy borrowing in the discourse. Even where specific educational policies of other countries are described, it is rarely in the context of applying pressure for policy reform. Why is this the case?

Externalization theory suggests, and other comparative education research has demonstrated, that references to foreign examples are key strategies of legitimization during times of contested domestic policy debate. Though these "selective description[s] and evaluative interpretation[s]" cannot be considered proper social scientific analysis, they can be used to great effect to shift both discourse and policy (Schreier and Martinez, 2004). But in the United States, despite the contentious implementation of

such policies as the Common Core State Standards during the late 2000s, references to PISA have not been used to legitimate domestic policy. In the case of the Common Core, this is in spite of the fact that PISA was specifically referenced in the creation of the CCSS, as an important international benchmark. Why the cold shoulder now that the standards are being implemented?

In the next section I offer some possible explanations about why PISA use has been used the way it has in American educational discourse. I begin by reviewing the theoretical premises of indicators and the mechanisms through which they work. Then I discuss some of the politics that might influence the decision to use or not PISA data. Finally, I propose that the discourse itself has been rhetorically weak, and has not been a useful point of leverage for disrupting the status quo and resetting the education agenda.

PISA as Ranked Data

Despite the growth of PISA in US discourse over time, there are few calls in the US for any education policy reform linked to PISA. This is surprising since rankings are theorized to alter the ways in which people understand information, thereby provoking reactions that lead to changes in behavior (Espeland and Sauder, 2007).⁷¹ In addition, PISA data are simplistic, reductive and potentially easy to use. In an era where science and numbers are proxies for impartiality and fairness (Desrosières, 1998), PISA

⁷¹ The glorification of Finland has, however, resulted in American education experts making pilgrimage to Finland to observe firsthand what the Finnish education system does. It has also put a major edutourism burden on Finland, where the Ministry of Education hosted at least 100 delegations per year from 2005-2011 (Loveless, 2014).

indicators provide an ‘objective’ look at how countries perform educationally. In an era where accountability has become the watchword in education, a test that allows policy makers and the public to see how well they compare to the international competition is compelling.

As with other kinds of rankings, PISA gains much of its power through a process of commensuration, whereby individual country units are made comparable. Kelley and Simmons (2015) offer three mechanisms through which performance indicators can affect policy: through domestic politics, through naming and shaming, and through markets. However PISA has not functioned quite along these lines. Taking the three mechanisms in reverse, PISA results have not moved markets as a credit rating would. For one thing, the OECD is not responsible for the funding of national education systems. There is no punishment for low scores, nor any reward (like a higher credit rating) for doing well. The market leverage that other kinds of performance indicators can have is missing from PISA.

Additionally, the degree to which shaming and blaming can occur is limited due to the fact that PISA scores do not provide the same normative evaluation as other indicators. Though PISA has set a standard of educational excellence, it is not imbued with the same normative meaning as indicators that monitor country performance such as human trafficking or other human rights violations. There is not a clear moral or ethical angle to PISA data, weakening its suitability for a naming-and-shaming strategy. Neither other

states nor the OECD have used PISA results to pressure educational change from the outside. To date PISA has worked most strongly as by affecting domestic politics, which I turn to next.

PISA and Politics

Politically there may be several reasons the PISA has been used, or not, in American discourse in the way that it has. The fact that PISA is only administered to approximately 5000 American students every three years, and is the product of an international organization based in Paris, means that the direct effect of PISA on American students is minimal. It may be that the distance between most students and PISA is large enough that PISA is not primed in the public mind to be an compelling, concerning or important reference. If neither elites nor the public has enough prior knowledge about international assessments to draw on (Chong and Druckman, 2007), references to PISA will be weak and out-of-context. It may be that domestic discourse is better served by inter-state comparisons, as from NAEP, than by referring to an assessment like PISA.

Secondly, it is unclear what PISA scores should be used for. The fundamental point that the OECD makes about PISA, that it is a tool for policymakers, is problematic when it comes to actually divining policy reform ideas from PISA scores. Precisely because PISA scores are not tied to any curriculum and are highly decontextualized, they cannot be used to great effect for anything but a proverbial horse race. An indicator system that is context free cannot be used for policy.

Third, it may be that using PISA data is not politically useful in the current educational climate. The lack of leverage that PISA is given may have to do with a kind of ‘scandalization’ fatigue. On this view, people are tired of a crisis narrative, and creating more crisis around PISA scores would not be politically useful. Given the growing opposition to the standardized testing that aligns to the Common Core, and the American accountability regime more generally, it may be that there is no room for more scandalization. In the minds of policy makers, PISA performance scores offer nothing new; they are no better or worse than they have historically been, and to use PISA for more scandalization may not be considered prudent.

Another view would see the use of PISA data as politically burdensome for other reasons. PISA could be used to aggravate both the political right and left, offering little political gain. On the left, the anti-accountability faction, “suburban soccer moms,” view PISA as one more arm of the accountability machine. On the right, an international assessment like PISA would be an overreach into states’ rights, as states are where most educational responsibility still lies. If this is the case, PISA would do more to anger than inspire the electorate.

PISA as Discourse

Finally, I return to the use of PISA in discourse to focus on how PISA is used as evidence in rhetoric. How a problem gets defined affects which actors pay attention, and how people relate specific problems to their own interests (Mintrom and Norman, 2009).

Scholarship has detailed some of the key characteristics of successful problem definition in public discourse, including causality (what is the cause of America's disappointing PISA scores?); the assignment of blame (whose fault is the current state of affairs?) and responsibility (what is the solution to the problem? who is responsible for the solution?); and consequences (what is at stake if the problem is not fixed?). Coupled with the use of quantitative indicators, simplistic story-telling, metaphors and appropriate symbols or values, these characteristics are more likely to construct a powerful discourse with strong problem definitions (Gusfield, 1981; Stone 2001; Schmidt, 2008; Mehta, 2011).

Thus far, the American PISA discourse has not clearly defined a problem, assigned blame or made a persuasive case about consequences. It is not clear from the texts what exactly the 'problem' of average to below-average PISA scores is. The United States has never done especially well on PISA or other international assessments; it has always done middling at best. The numbers are such that invoking PISA performance does not allow elites to a call for a return to the 'good old days' or a break from the past.

The agenda setting stage of policy making requires disruption, as new ideas enter the fray or preexisting ideas take hold, or take hold in new ways. The rhetoric about PISA does not redefine educational problems or force a reconsideration of long-held beliefs about American education. Instead the PISA discourse aligns with existing values and supports prevailing understandings about the American education system, the narrative of decline that has been in place since the release of *A Nation at Risk*. Like Japan in the 1980s,

which was viewed as an economic threat and rival and thus became a reference society in educational discourse, China has taken become a similar villain in contemporary crisis language. While some of the imagery is strong, it does not disrupt the zeitgeist of mediocrity that has defined American education for the past thirty years. This narrative does not force a reconsideration of the ways in which US education has been represented, particularly in the US press, for decades.

Agreement on meaning. There is broad agreement across literatures and sources on the meaning of PISA scores. The think tanks and the media, regardless of political orientation, use the same frames. The discourse does little to disrupt popular thinking, since the frames of the PISA discourse are already deeply embedded in the US psyche. The primary frame is linking education to economic development, at both the state and individual levels. Since the 1980s, the quasi-market language of accountability, choice, consumer management and competition has been the basis of discourse of US education policy. Economic competitiveness is the *raison d'être* for schooling, and economic logic is used to justify a range of reforms. This approach includes authoritative language that links economic competitiveness to reform in education, even if there is mixed evidence of the “education spending in, economic growth out” that some reformers describe (Coffield, 2011, 133).

Conflicting reasoning. This is seen in the conflicting explanations of success attributed to top scorers. In Finland, teachers and equity are important variables; in

China, it is government education policy. How should policy makers determine the merits of each approach and decide which practices are best? The same problem is evident in explanations of culture. Finnish success is driven by a communitarian culture and by social democracy that provides a deep social safety net for its citizens. In China, the literature points to an “Asian” outlook that upholds a Confucian reverence for education as the culture of success. But these cultural explanations are a conundrum for policy makers. Culture is contingent, and not easily replicable. PISA proponents see cultural differences as a threat to cross-national borrowing. While policies are seen as readily changeable – at least, according to the various reports about what can be learned from ‘strong’ and ‘best’ performing systems – cultural explanations are seen as pessimistic and even defeatist. On this view, “cultural and historical explanations for the success of education systems cannot be used to justify reforms in other nations, whereas pointing to specific policy settings as the cause of success can provide governments with leverage for internal reform agendas” (Sellar and Lingard, 2013b, 14).

These conflicting explanations of success are potentially problematic for US policy makers since they offer divergent claims about successful practice. That the references to foreign examples can be all things to all people provides only weak policy guidance. On the other hand, this juxtaposition also potentially creates a space for debating US education reform. By employing PISA results generally, and Finland and China specifically, different actors can offer a range of claims about ‘what works’ in education.

Role of the state. Additionally, in much of the discourse around reference societies success is predicated on a strong interventionist role for the state. In Finland, the historical contingencies of Soviet domination and a largely agrarian society led the state to commit to the development of a system of comprehensive schooling over decades of successive governments (Simola, 2005; Porter-Magee, 2012). More recently in China, government commitment to ‘national honor’ and being a world leader in education is credited with pushing additional resources and pressure to Shanghainese students to achieve top scores (Finn, 2010; Loveless, 2014). PISA performance in these narratives is derivative of strong and specific government action. This discursive framing may offer little of interest to many reformers in the United States, where the state is often problematized in educational discourse. The state-as-roadblock-to-quality-education has been a prominent refrain of many American education reformers of the past three decades. Seen most obviously in the movement towards public schooling through mechanisms of choice, the dominant view towards the state in US education discourse has been one of distrust.

Areas for Future Work

My research adds to a nascent body of scholarship that looks at international assessment studies as a field of academic research. Comparative case work (e.g., Takayama, 2010; Takayama, Waldow and Sung, 2013) has begun to show the nuanced effects of the PISA discourse across national settings, but more systematic work needs to be done to specify the processes and networks through which ideas about ILSAs travel and become accepted

discursive tropes. Future areas for research could include empirical work in such areas as how ideas travel amongst educational elites, the role of public opinion and elite responsiveness in discourse formation, and the roles that culture and national identity play in policy reception. Future work might also look more closely and comparatively at the mechanisms of reception; at what moments in the policy process is PISA most useful? Why does reception shift from one type of reaction to another? Under what conditions are PISA data most successful in influencing policy debate?

Indifference as a Response

I conclude with a note of caution about how to define the US reception of PISA. I have shown that after a limited period of crisis talk, PISA reception in the US has been rather indifferent. But this is perhaps not the best characterization. As Gusfield (1981) observed, indifference is not a passive response, but an active choice. Rather than indifferent, perhaps we should say that actors are not activating PISA at the moment. They have not come together to form new strategic coalitions or advocacy networks (e.g., Sabatier and Jenkins-Smith, 1993), nor have education elites used the empirical data of PISA to great effect. But ignoring PISA has its own logic as well; in this sense, even the lack of conflict may be evidence of a conflict because it is contributing to ‘what everyone knows’ (Gusfield, 1981). “Indifference” is not a passive response, but a determination that PISA data is not useful at the moment.

As Kingdon and others have discussed, the policy cycle is not a continuous process, but can proceed unpredictably. It is clear that PISA is not seen as useful *right now*, but that does not mean it will not be used as evidence in the future. As the political climate shifts towards identifying solutions to problems like the implementation of the Common Core, and backlash against its testing regime, time may come when policy entrepreneurs or other agents may pull PISA out of the ‘primordial soup’ of ideas to help construct clear problems and solutions in the education arena. PISA may prove useful in such future deliberations.

Conclusion

What is clear is that the idea of PISA has become rapidly resonant in educational discourse. Though it has yet to have a measurable policy impact in the United States – nothing programmatic has come of PISA – PISA has become a part of the rhetorical toolkit of education elite. Although the PISA narrative in the United States has not led to new institutionalized forms of action, it has impacted elite discourse. Elites across institutions and across the political spectrum are using PISA as evidence in educational policy discourse. This suggests that PISA has ideational qualities that resonate with elite policy actors.

In this dissertation, I highlight the subtle ways in which ideas come to be defined and accepted in the public arena. I show that problem definition is a contested process amongst actors with varying levels of power and persuasiveness. Drawing on the

theoretical framework of externalization, my findings support the contention that much educational borrowing taking place between countries is at the discursive, rather than the programmatic level. But these discursive adoptions should be taken seriously. The ideas that gain purchase are those that will influence which problem definition takes hold; the ways in which problems are defined influences which actors can act to provide solutions and which actions can be taken. Discourse shapes and constrains policy talk and policy action.

This dissertation is a first step in understanding PISA reception in the United States. I have aimed to elucidate the ways in which actors, agency and ideas interact to develop discourse around PISA in the United States. In unpacking the relationships between agency, incentives and structure, my research offers some insight into the how educational problems come to be defined in the public arena. My research suggests that norms, geopolitics, culture and timing are part of the judicious mix that determine whether and how certain educational memes take hold domestically. Articulating the role of academic, think tank and media literature in creating the discourse allows for a more detailed understanding of how American elites problematize education. It also suggests that a critical ear is necessary to understand what we talk about when we talk about PISA.

References

- AERA. (2015). About AERA. Retrieved from <http://www.aera.net/AboutAERA/tabid/10062/Default.aspx>.
- Afonso, N. and Costa, E. (2009). Use and circulation of OECD's "Programme for International Student Assessment" (PISA) in Portugal. *KnowandPol: Supra-national instruments*. Retrieved from www.knowandpol.eu.
- Alliance for Audited Media. (2013). Top 25 US Newspapers for March 2013. Retrieved from www.auditedmedia.com/news/blog/top-25-us-newspapers-for-march-2013.aspx.
- Amermuller, A. (2005). *Poor Background or Low Returns? Why Immigrant Students in Germany Perform so Poorly in PISA*. ZEW Discussion Papers, No. 05-18. Leibniz-Information Centre for Economics. Retrieved from <http://hdl.handle.net/10419/24111>.

- American Educational Research Association. (2011). Summary of AERA journal operations, 2010. *Educational Researcher*, 40(3), 128-129. DOI: 10.3102/0013189X11410162.
- Bajoni, I., Berényi, E., Neumann, E., and Vida, J., The reception of PISA in Hungary. *KnowandPol: Supra-national instruments*. Retrieved from www.knowandpol.eu.
- Baumgartner, F. and Jones, B. (1993). *Agendas and instability in American politics*. Chicago: University of Chicago Press.
- Béland, D., and Orenstein, M.A. (2012). International organizations as policy actors: An ideational approach. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2165514.
- Benford, R.D. and Snow, D.A. (2000). Framing processes and social movements: An overview and assessment. *Annual Review of Sociology*, 26, 611-639.
- Bereday, G.Z.F. (1964). Sir Michael Sadler's "Study of foreign systems of education." *Comparative Education Review*, 7(3), 307-314. <http://www.jstor.org/stable/118711>
- Berliner, D. and Biddle, B. (1995). *The manufactured crisis: Myths, fraud and the attack on America's public schools*. White Plains, NY: Longman.
- Bieber, T. and Martens, K. (2011). The OECD PISA study as soft power in education? Lessons from Switzerland and the US. *European Journal of Education*, 46(1), 101-116.
- Bjork, C. and Tsuneyoshi, R. (2005). Education reform in Japan: Competing visions for the future. *Phi Delta Kappan*, April 2005. 619-626.
- Breakspear, S. (2012). *The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance*. (OECD Education Working Papers, No. 71). OECD Publishing. Retrieved from <http://dx.doi.org/10.1787/5k9fdfqffr28-en>.
- Brown, E. (2015, January 25). What happens when the Common Core become less...common? *The Washington Post*. Retrieved from http://www.washingtonpost.com/local/education/what-happens-when-the-common-core-becomes-less-common/2015/01/25/33b8eb58-a2bf-11e4-b146-577832eafcb4_story.html.

- Brookings Institution. (2015). Brown Center on Education Policy. Retrieved from <http://www.brookings.edu/about/centers/brown>.
- Bybee, R.W. (1997). “*The Sputnik Era: Why is this educational reform different from all other reforms?*” Paper from a symposium hosted by the Center for Science, Mathematics and Engineering Education, National Academies of Science.
- Campbell, J. (2002). Ideas, politics and public policy. *Annual Review of Sociology*, 28, 21-38. DOI:10.1146/annurev.soc.28.110601.141111.
- Carney, S., Rappleye, J., and Silova, I. (2012). Between faith and science: World culture theory and comparative education. *Comparative Education Review*, 56(3), 366-393. <http://www.jstor.org/stable/10.1086/665708>.
- Carnoy, M. and Rothstein, R. (2013a). What do international tests really show about US student performance? Washington, DC: Economic Policy Institute.
- Carnoy, M. and Rothstein, R. (2013b). Response from Martin Carnoy and Richard Rothstein to OECD/PISA Comments (by Andreas Schleicher, OECD Deputy Director for Education and Special Advisor on Education Policy to the OECD’s Secretary-General, January 14, 2013) regarding our report “What do international tests really show about American student performance?” Washington, DC: Economic Policy Institute.
- Carvalho, M. The fabrications and travels of a knowledge-policy instrument. *European Education Research Journal*, 12(2), 172-188.
- Chong, D. and Druckman, J.N. (2007). A theory of framing and opinion formation in competitive elite environments. *Journal of Communication*, 57, 99-118. DOI:10.1111/j.1460-2466.2006.00331.x
- Chong, D. and Druckman, J.N. (2011). Public-elite interactions: puzzles in search of researchers. In *The Oxford Handbook of American Public Opinion and The Media*. Edwards, G., Jacobs, L. and Shapiro, R. (eds.) From Oxford Handbooks Online. DOI: 10.1093/oxfordhb/9780199545636.003.0011.
- Chubb, J. and Moe, T. (1990). *Politics, Markets and America’s Schools*. Washington, DC: The Brookings Institute.
- Coffield, F. (2012). Why the McKinsey report will not improve school systems. *Journal of Education Policy*, 27(1): 131-149.
- Common Core State Standards Initiative Standard Setting Criteria. (n.d.) Retrieved from <http://www.corestandards.org/assets/Criteria.pdf>.

- Coppedge, M. (1999). Thickening thin concepts and theories: combining large-N and small in comparative politics. *Comparative Politics*, 31(4), 465-476. Stable URL: <http://www.jstor.org/stable/42224>.
- Cresswell, J.W. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. (2nd edition.). Thousand Oaks, CA: Sage Publications.
- Cummings, W. K. (1989). The American perception of Japanese education. *Comparative Education*, (25), 293–302.
- Cummings, W.K. (2007). The American Perception of Japanese Education. In Crossley, M., Broadfoot, P. and Schweisfurth, M. (Eds.) *Changing Educational Contexts, Issues and Identities: 40 Years of Comparative Education*. (95-107). New York: Routledge.
- Dahl, R. (1958). A critique of the ruling elite model. *The American Political Science Review*, 52(2), 463-469. <http://www.jstor.org/stable/1952327>.
- Davis, K., Kingsbury, B., and Merry, S. (2012). Introduction: global governance by indicators. In Davis, K., Fischer, A., Kingsbury, B., and Merry, S. (Eds.) *Governance by Indicators: Global Power through Quantification and Rankings*. Oxford Scholarship Online. DOI: 10.1093/acprof:oso/9780199658244.001.0001.
- DeBray, E.H. (2006). *Politics, ideology and education: Federal policy during the Clinton and Bush administrations*. New York: Teachers College Press.
- DeBray-Pelot, E. and McGuinn, P. (2009). The new politics of education: Analyzing the federal education policy landscape in the post-NCLB era. *Educational Policy*, 23(1); 15-42. DOI: 10.1177/0895904808328524.
- DeCoker, G. (2013). Enduring issues, new contexts. In DeCoker, G. and Bjork, C. (Eds). *Japanese Education in an Era of Globalization: Culture, politics and equity*. New York: Teachers College Press.
- Desrosières, A. (1998). *The politics of large numbers: A history of statistical reasoning*. Cambridge, MA: Harvard University Press.
- Dixon, R., Arndt, C., Mullers, M., Vakkuri, J., Engblom-Pelkkala, K., and Hood, C. (2012). A lever for improvement of a magnet for blame? Press and political responses to international educational rankings in four EU countries. *Public Administration*, 91(2), 282-505. DOI: 10.1111/padm.12013.

- Dobbins, M. & Martens, K. (2012). Towards an education approach à la Finlandaise? French education policy after PISA. *Journal of Education Policy*, 27(1), 23-43. DOI: 10.1080/0305006032000162002.
- Döbert, H. (2007). Germany. In Horner, W., Döbert, H., von Kopp, B., and Mitter, W. (Eds.). *The Education Systems of Europe*. Dordrecht: Springer.
- Duncan, A. (2010, December 7). Remarks at eh OECD's Release of the Programme for International Student Assessment (PISA) 2009 Results. Retrieved from <http://www.ed.gov/news/speeches/secretary-arne-duncans-remarks-oecd-release-program-international-student-assessment-pisa-2009-results>.
- Engel, L.C. and Frizzell, M. (2015). Competitive comparison and PISA bragging rights: sub-national uses of the OECD's PISA in Canada and the USA, *Discourse: Studies in the Cultural Politics of Education*, (1-18). DOI: 10.1080/01596306.2015.1017446.
- Ertl, H. (2006). Educational standards and the changing discourse on education: The reception and consequences of the PISA study in Germany. *Oxford Review of Education*, 32(5), 619-634.
- Espeland, W.N. and Sauder, M. (2007). "Rankings and reactivity: How public measures recreate social worlds." *American Journal of Sociology*, 113(1), 1-40.
- Espeland, W.N. and Stevens, M.L. (1998). Commensuration as a social process. *Annual Review of Sociology*, 24, 313-343.
- Farrell, H. and Drezner, D. (2008). The power and politics of blogs. *Public Choice* (134) ½, 15-30. www.jstor.org/stable/27698208.
- Feniger, Y. and Lefstein, A. (2014). How not to reason with PISA data: an ironic investigation. *Journal of Education Policy*, DOI: 10.1080/02680939.2014.892156.
- Feniger, Y., Livneh, I., & Yogev, A. (2012). Globalisation and the politics of international tests: The case of Israel. *Comparative Education*, 48(1), 323-335.
- Ferree, M.M., Gamson, W., Gerhards, J., and Rucht, D. (2002). *Shaping abortion discourse: Democracy and the public sphere in Germany and the United States*. New York: Cambridge University Press.
- Feuer, M. (2012). No country left behind: Rhetoric and reality of international large-scale assessment. The 13th William Angoff Memorial Lecture. Washington, DC: Educational Testing Service.

- Fischman, G. and Tefera, A. (2014). If the research is not used, does it exist? *Teachers College Record*, Commentary. Retrieved from <http://www.tcrecord.org>.
- Fiss, P.C. and Hirsch, P.M. (2005). The discourse of globalization. *American Sociological Review*, 70(1), 29-52.
- Fladmoe, A. (2011). Education in the news and in the mind: PISA, news media and public opinion in Norway, Sweden and Finland. *Nordicom Review*, 32(2), 99-116.
- Frank, D. J., Hironaka, A., & Schofer, E. (2000). The nation-state and the natural environment over the twentieth century. *American Sociological Review*, 65(1), 96-116.
- Gamson, W.A. (1992). *Talking Politics*. Cambridge, UK: Cambridge University Press.
- Gamson, W.A. and Modigliani, A. (1989). Media discourse and public opinion on nuclear power: a constructionist approach. *American Journal of Sociology*, 95(1), 1-37.
- Gellner, W. (1995). The politics of policy “political think tanks” and their markets in the US institutional environment. *Presidential Studies Quarterly*, 25(3), 497-510. www.jstor.org/stable/27551464.
- Goldhaber, D. & Brewer, D. (2008). What gets studied and why: examining the incentives that drive education research. In Hess, F. (Ed.). *When Research Matters: How Scholarship Influences Education Policy*. Cambridge, MA: Harvard Education Press.
- Gordon, J. and LeTendre, G. (2010). New policy context for schooling in Japan. In Gordon, J., Fujita, H., Kariya, T. LeTendre, G. (Eds). *Challenges to Japanese Education: Economics, reform, human rights*. New York: Teachers College Press.
- Guthrie, J.W. and Springer, M.G. (2004). “A Nation at Risk” revisited: Did “wrong” reasoning result in “right” results? At what cost? *Peabody Journal of Education*, 79(1), 7-35. <http://www.jstor.org/stable/1493062>.
- Gusfield, J. (1981). *The Culture of Public Problems*. Chicago: University of Chicago Press.
- Grek, S. (2009). Governing by numbers: the PISA “effect” in Europe. *Journal of Education Policy*, 24(1), 23-37. DOI: 10.1080/02680930802412669.

- Grek, S., M. Lawn, B. Lingard, J. Ozga, R. Rinne, C. Segerholm, et al. (2009). "National Policy Brokering and the Construction of the European Education Space in England, Sweden, Finland and Scotland. *Comparative Education*, 45 (1), 5–21.
- Haas, E. (1990). *When knowledge is power: three models of change in international organizations*. Berkeley, CA: University of California Press.
- Heckman, J. and Kautz, T. (2013). Fostering and measuring skills: Interventions that improve character and cognition. *NBER Working Paper Series*, Working Paper 19656. Retrieved from <http://www.nber.org/papers/w19656>.
- Henig, J.R. (2008a). The Evolving Relationship between Researchers and Public Policy. *Phi Delta Kappan*. 89(5), 357-360. Retrieved from http://www.pdkmembers.org/members_online/publications/Archive/pdf/k0801hen.pdf.
- Henig, J.R. (2008b). *Spin Cycle: How research is used in policy debates: The case of charter schools*. New York, NY: Russell Sage Foundation and the Century Foundation.
- Henry, M., Lingard, B., Rizvi, F. and Taylor, S. (2001). *The OECD, Globalization and Education Policy*. Oxford: Pergamon Press.
- Hess, F. (Ed). (2008). *When Research Matters: How Scholarship Influences Education Policy*. Cambridge, MA: Harvard Education Press.
- Heyneman, S.P. (1993). "Quantity, Quality, and Source." *Comparative Education Review*, 37 (4), 372-388.
- Howie, S., & Plomp, T. (2005). International comparative studies of education and large-scale change. In *International Handbook of Educational Policy*. (75-99). Rotterdam: Springer Netherlands.
- Ikenson, D. (2013, January 29). Soured US-China relationship approaches inflection point. Retrieved from www.forbes.com.
- Kameda, M. (2013). Nation's kids top fields in PISA test. Retrieved from <http://www.japantimes.co.jp/news/2013/12/03/national/nations-kids-top-fields-in-pisa-test/>.
- Kamens, D. H. (2013). Globalization and the emergence of an audit culture. In Meyer, H-D. and Benavot, A. (Eds.) *PISA, power and policy: The emergence of global educational governance*. (p. 117-139). Oxford, UK: Symposium Books.

- Kamens, D. H., and Benavot, A. (2011). National, regional and international learning assessment: trends among developing countries, 1960-2009. *Globalisation, Societies and Education*, 9(2), 285-300. DOI: 10.1080/14767724.2011.577337.
- Kamens, D. and McNeely, C. (2010). Globalization and the growth of international educational testing and national assessment, *Comparative Education Review*, 54(1), 5-25.
- Kandel, I.L. (1959). The methodology of comparative education. *International Review of Education*, 5(3), 270-280. <http://www.jstor.org/stable/3441523>.
- Kelley, J.G. and Simmons, B.A. (2015). Politics by numbers: indicators as social pressure in international relations. *American Journal of Political Science*, 59(1), 55-70. DOI: 10.1111/ajps.12119.
- Kingdon, J. (2011). *Agendas, alternatives and public policies*. (Updated 2nd Edition). New York: Longman.
- Koretz, D. (2008). *Measuring up: What educational testing really tells us*. Cambridge, MA: Harvard University Press.
- Kristof, N. (2014, February 16). Professors, we need you! New York Times. SR11. Retrieved from http://www.nytimes.com/2014/02/16/opinion/sunday/kristof-professors-we-need-you.html?_r=0.
- Labaree, D. (2014). Let's measure what no one teaches: PISA, NCLB, and the shrinking aims of education. *Teachers College Record*, 116, 1-14.
- Lehmann, R. (2011). "The impact of IEA on Educational Policy Making in Germany." In IEA 1958-2008: 50 years of experiences and memories. Volume I. Papanastasiou, C., Plomp, T. and Papanastasiou, E.C. (Eds). Nicosia, Cyprus: Cultural Center of the Kykkos Monastery. P. 411-430.
- LeTendre, G. and Baker, D. (1999). International comparisons and educational research policy. In LeTendre, G. (Ed). *Competitor or ally? Japan's role in American educational debates*. (123-139). New York: Falmer Press.
- Levin, B. (2004, October 17). Making research matter more. *Education Policy Analysis Archives*, 12(56). Retrieved from <http://epaa.asu.edu/epaa/v12n56/>.
- Levin, H.M. (2012). More than just test scores. *Prospects: Quarterly review of comparative education*. 137(1): 269-284.

- Lingard, B. & Grek, S. (2007). The OECD, indicators and PISA: An exploration of events and theoretical perspectives. (Working Paper 2). *ESRC/ESF Research Project on Fabricating Quality in Education*. Retrieved from www.ces.ed.ac.uk/PDF%20Files/FabQ_WP2.pdf.
- Lingard, B. and Rawolle, S. (2011). New scalar politics: Implications for education policy. *Comparative Education*, 47(4), 489-502. DOI: 10.1080/03050068.2011.55594.
- Lockheed, M. (2013). Causes and consequences of international assessments in developing countries. In Mayer, H-D. and Benavot, A. (Eds.) (2013). *PISA, power and policy: The emergence of global educational governance*. (p. 163-184). Oxford: Symposium Books.
- Loveless, T. (2009). 2008 Brown Center Report on Education. Retrieved from <http://www.brookings.edu/research/books/2009/the2008browncenterreportonamericaneducation>.
- Loveless, T. (2012). Misinterpreting International Test Scores. In 2012 Brown Center Report on American Education. Retrieved from <http://www.brookings.edu/research/reports/2012/02/16-brown-education>.
- Loveless, T. (2014a). Lessons from the PISA Shanghai controversy. In 2014 Brown Center Report on American Education. Retrieved from <http://www.brookings.edu/research/reports/2014/03/18-pisa-shanghai-loveless>.
- Loveless, T. (2014b, November 20). "The perils of edutoursim." The Brown Center Chalkboard. Retrieved from www.brookings.edu/research/papers/2014/11/20-chalkboard-edutourism-loveless.
- Manna, P. (2006). *School's in: Federalism and the national education agenda*. Washington, DC: Georgetown University Press.
- Manna, P. (2011). *Collision course: Federal education policy meets state and local realities*. Washington, DC: CQ Press.
- Martens, K. (2008). Why is there no PISA shock in the US? A comparison of German and American education policy. *American German Institute for Contemporary German Studies at Johns Hopkins University*. Retrieved from <http://www.aicgs.org>.
- Martens, K. and Niemann, D. (2010). Governance by comparison: How ratings and rankings impact national policymaking in education. *TranState Working Papers*, 139. Retrieved from <http://www.staatlichkeit.uni-bremen.de>.

- McCombs, M. and Shaw, D.L. (1972). The agenda setting function of mass media. *Public Opinion Quarterly*, 36(2), 176-187.
- McDonald, L. (2014). Think tanks and the media: How the conservative movement gained entry into the education policy arena. *Educational Policy*, 28(6), 845-880. DOI: 10.1177/0895904813492372.
- McGann, J. (2013). *2012 Global Go To Think Tanks Report and Policy Advice*. Think Tanks and Civil Societies Program, University of Pennsylvania. Downloaded from <http://www.gotothinktank.com>.
- McGuinn, P. (2005). The national schoolmarm: No Child Left Behind and the new education federalism. *Publius*, 41-68.
- McGuinn, P. (2006). *No Child Left Behind and the Transformation of Federal Education Policy, 1965-2005*. Lawrence, KS: University Press of Kansas.
- Mehta, J.D. (2006). *The transformation of American educational policy, 1980-2001: Ideas and the rise of accountability politics*. Unpublished doctoral dissertation. ProQuest Dissertations and Theses Full Text.
- Mehta, J. (2011). The varied role of ideas in politics: From “whether” to “how.” In Béland, D. & Cox, R.H. (Eds.). *Ideas and Politics in Social Research*. Oxford: Oxford University Press. 23-46.
- Mehta, J. (2013). How paradigms create politics: The transformation of American educational policy, 1980-2001. *American Educational Research Journal*, 50 (20), 285-324.
- Meyer, H-D. and Benavot, A. (Eds.). (2013). *PISA, power and policy: The emergence of global educational governance*. Oxford: Symposium Books.
- Miles, M. and Huberman, A.M. (1998). *Qualitative data analysis: An expanded sourcebook*. 2nd Edition. Washington, DC: Sage Publications.
- Mintrom, M. and Norman, P. (2009). Policy entrepreneurship and policy change. *The Policy Studies Journal*, 37(4), 649-667. DOI: 10.1111/j.1541-0072.2009.00329.x
- Morgan, S.L. and Taylor-Poppe, E.S. (2012). The consequences of international comparisons for public support of K-12 education: Evidence from a national survey experiment. *Educational Researcher*, 41(7), 262-268. DOI: 10.3102/0013189X12455217.

- National Center for Education Statistics, US Department of Education (n.d.) An introduction to NAEP: National assessment of educational progress. Retrieved from <http://nces.ed.gov/nationsreportcard/pdf/parents/2010468.pdf>.
- National Center for Education Statistics, US Department of Education. (2010). *Teacher Attrition and Mobility: Results from the 2008–09 Teacher Follow-up Survey* (NCES 2010-353). Retrieved from <http://nces.ed.gov/fastfacts/display.asp?id=28>.
- 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, DC: The Commission. Retrieved from <https://www2.ed.gov/pubs/NatAtRisk/risk.html>.
- National Research Council. (2012). *Using science as evidence in public policy*. Committee on the Use of Social Science Knowledge in Public Policy, Prewitt, K., Schwandt, T. A., & Straf, M. L. (Eds). Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Noah, H. J. (1985). “Comparative education: methods.” In T. Husen and TN. Postlethwaite (eds.), *The International Encyclopedia of Education*. 2nd Edition, New York: Pergamon.
- Nóvoa, A. and Yariv-Mashal T. (2003) ‘Comparative research in education: a mode of governance or a historical journey?’ *Comparative Education*, 39(4), 423-438.
- Obama, B. (2010, December 6). Speech at Forsyth Technical Community College, Winston-Salem, North Carolina. Retrieved from www.whitehouse.gov.
- O’Brien, P. and Pizmony-Levy, O. (in press). Going public, gaining credibility: student perceptions of publicly engaged scholars. *Sociological Perspectives*.
- OECD. (2001). *Knowledge and skills for life: First results from PISA 2000*. Paris: OECD. Retrieved from www.oecd.org/edu/school/programmeforinternationalstudentsassessmentpisa/22691596.pdf.
- OECD. (2003). PISA 2003 technical report. Paris: OECD. Retrieved from <http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/35188570.pdf>.
- OECD. (2014a). PISA tri-fold brochure. Retrieved from <http://www.oecd.org/pisa/aboutpisa/PISA-trifold-brochure-2014.pdf>.

- OECD. (2014b). Response to points raised in Hans-Dieter Meyer's Open Letter. Retrieved from <http://www.oecd.org/pisa/aboutpisa/OECD-response-to-Heinz-Dieter-Meyer-Open-Letter.pdf>.
- Olafsdottir, S. (2007). *Medicalizing mental health: a comparative view of the public, private and professional construction of mental illness*. Unpublished doctoral dissertation. Indiana University. ProQuest Dissertations and Theses Full Text.
- Owens, T. L. (2013). Thinking beyond league tables: a review of key PISA research questions. In Meyer, H-D. and Benavot, A. (Eds.) *PISA, power and policy: The emergence of global educational governance*. (p. 27-49). Oxford, UK: Symposium Books.
- Ozga, J. (2012). "Comparison as a governing technology: The case of PISA. *Research Intelligence*, 119, 18-19.
- Papadopoulos, G. (2011). The OECD approach to education in retrospect 1960-1990. *European Journal of Education*, 46(19), 85-86.
- Park, H. and Lee, Y-J. (2013) Growing educational inequality in Japan during the 2000s. In DeCoker, G. and Bjork, C. (Eds.) *Japanese Education in an Era of Globalization: Culture, politics and equity*. (131-146). New York: Teachers College Press.
- Pizmony-Levy, O. (2013). *Testing for all: The emergence and development of international assessment of student achievement, 1958—2012*. Unpublished doctoral dissertation. Indiana University, ProQuest Dissertations and Theses Full Text.
- Pizmony-Levy, O. (2014). Back to the Future on International Assessments. In Chatterji, M. (Guest Ed.). On the merits of, and myths about, international assessments, *Quality Assurance in Education*, 22(4), 319 – 338. DOI: 10.1108/QAE-07-2014-0035.
- Pongratz, L. (2006) Voluntary self-control: Education reform as a governmental strategy. *Educational Philosophy and Theory*, 38 (4). 471-484.
- Pons, X. (2012). Going beyond the 'PISA shock' discourse: an analysis of the cognitive reception of PISA in six European countries, 2001-2008. *European Educational Research Journal*, 11(2),-226.
- Pons, X. and van Zanten, A. (2007). Knowledge circulation, regulation and governance. *KnowandPol Literature Review*, 10-38.

- Postlethwaite, T.N. (1987). "Comparative educational achievement research: can it be improved?" *Comparative Education Review*, 31(1). Special Issue on the Second IEA Study. 150-158.
- Ragin, C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley, CA: University of California Press.
- Rapplee, J. (2012). Reimagining attraction and 'borrowing' in education: Introducing and political production model. In Steiner-Khamsi, G. and Waldow, F. (Eds). *World Yearbook of Education: Policy borrowing and lending in education*. New York: Routledge.
- Rich, A. (2004). *Think Tanks, Public Policy and the Politics of Expertise*. New York, NY: Cambridge University Press.
- Rich, A. and Weaver, R.K. (2000). Think tanks and the media. *The Harvard International Journal of Press and Politics*, 5(4), 81-103.
- Ringarp, J. and Rothland, M. (2010). Is the grass always greener? The effect of the PISA results on education debates in Sweden and Germany. *European Educational Research Journal*, 9(3), 422-430. DOI: [10.2304/eeerj.2010.9.3.422](https://doi.org/10.2304/eeerj.2010.9.3.422).
- Róstaş, Z., Kósa, I., Bodó, J., Kiss, A., Fejes, I. (2009). Use and Circulation of PISA in a Romanian context. In *KnowandPol: Supra-national instruments*. Retrieved from www.knowandpol.eu.
- Rothstein, H. & Hopewell, S. (2009). Grey literature. In Cooper, H., Hedges, L., and Valentine, J. (Eds.). *Handbook of Research Synthesis and Meta-Analysis*. 2nd Edition (pp. 104-122). New York, NY: Russell Sage Foundation.
- Rubner, J. (2006). How can a country manage the impact of 'poor' cross-national research results? In Ross, K. and Genevois, I. (Eds). *Cross-national studies of the quality of education: planning their design and managing their impact*. Paris: UNESCO.
- Sabatier, P.W. and Jenkins-Smith, H.C. (1993). *Policy change and learning: an advocacy coalition approach*. Boulder, CO: Westview Press.
- Sambanis, N. (2004). Using case studies to expand economic models of civil war. *Perspectives on Politics* 2 (2): 259-280. Retrieved from <http://pantheon.yale.edu/~ns237/methods.html>.

- Sellar, S. and Lingard, B. (2013a) Looking East: Shanghai, PISA 2009 and the reconstitution of reference societies in the global education policy field, *Comparative Education*, 49(4), 464-485, DOI: 10.1080/03050068.2013.770943.
- Sellar, S. and Lingard, B. (2013b). PISA and the expanding role of the OECD in global educational governance. In Meyer, H-D. and Benavot, A. (Eds.) *PISA, power and policy: The emergence of global educational governance*. (p. 185-206). Oxford, UK: Symposium Books.
- Schriewer, J. (1990). The method of comparison and the need for externalization: Methodological criteria and sociological concepts. In Schriewer, J. and Holmes, B. (Eds). *Theories and methods in comparative education*. (3rd Edition), (25-38). Frankfurt: Lang.
- Schriewer, J. & Martinez, C. (2004). Constructions of internationality in education. In Steiner-Khamsi, G. (Ed.). *The Global Politics of Educational Borrowing and Lending*. New York, NY: Teachers College Press.
- Schmidt, V. (2008). Discursive institutionalism: The explanatory power of ideas and discourse. *Annual Review of Political Science*, 11, 303-326. DOI: 10.1146/annurev.polisci.11.060606.135342.
- Schmidt, W., Hoang, R., and Shakrani, S. (2009). *International lessons about national standards*. Washington, DC: The Thomas B. Fordham Institute.
- Silova, I. (2006). *From symbols of occupation to symbols of multiculturalism: Re-conceptualizing minority education in post-Soviet Latvia*. Unpublished doctoral dissertation. ProQuest Dissertations and Theses Full Text.
- Smithers, A. (2004). England's education: What can be learned by comparing countries? Centre for Education and Employment Research, University of Liverpool.
- Snyder, J. and Ballentine, K. (1996). Nationalism and the marketplace of ideas. *International Security*, 21(2), 5-40. <http://www.jstor.org/stable/2539069>.
- Stack, M. (2006). Testing, testing read all about it: Canadian press coverage of the PISA results. *Canadian Journal of Education*, 29(1), 49-69.
- Stack, M. (2007). Representing school success and failure: media coverage of international tests. *Policy Futures in Education*, 5(1), 100-110. DOI: 10.2304/pfie.2007.5.1.100.

- Stedman, L.C., (1994). Incomplete explanations: The case of US performance in the international assessments of education. *Educational Researcher*, 23(7), 24-32. DOI: 10.3102/0013189X023007024.
- Steiner-Khamsi, G. (2003). The politics of league tables. *Journal of Social Science Education*. Retrieved from <http://www.jsse.org/jsse/index.php/jsse/article/view/470/386>.
- Steiner-Khamsi, G. (Ed.). (2004). *The global politics of educational borrowing and lending*. New York: Teachers College Press.
- Steiner-Khamsi, G. (2010). The politics and economics of comparison. CIES 2012 Presidential Address. *Comparative Education Review*, 54(3), 323-342.
- Steiner-Khamsi, G. (2014). Cross-national policy borrowing: understanding reception and translation. *Asia Pacific Journal of Education*, 34(2), 153-167. DOI: 10.1080/02188791.2013.875649.
- Stone, D. (2001). *The policy paradox: The art of political decision making* (revised edition). New York: W.W. Norton.
- Swanson, C. B. and Barlage, J. (2006). *Influence – A study of the factors shaping education policy*. Editorial Projects in Education. Bethesda, MD.
- Takayama, K. (2007). A Nation at Risk crosses the Pacific: Transnational borrowing of the US crisis discourse in the debate on education reform in Japan. *Comparative Education Review*, 51, 423-446.
- Takayama, K. (2008). The politics of international league tables: PISA in Japan's achievement crisis debate. *Comparative Education*, 44(4), p. 387-407.
- Takayama, K. (2010). Politics of externalization in reflexive times: Reinventing Japanese education reform discourses through "Finnish PISA Success." *Comparative Education Review*, 55(1), 51-75.
- Takayama, K. (2012). Bringing political bite to educational transfer studies: Cultural politics of PISA and the OECD in Japanese education reform. In Steiner-Khamsi, G. and Waldow, F. (Eds). *World Yearbook of Education: Policy borrowing and lending in education*. New York: Routledge.
- Takayama, K., Waldow, F., Sung, Y.L. (2013) Finland Has it All? Examining the Media Accentuation of 'Finnish Education' in Australia, Germany and South Korea. *Research in Comparative and International Education*, 8(3), 307-325. DOI: 10.2304/rcie.2013.8.3.307.

- Takayama, K., Waldow, F. and Sung, Y.L. (2014). Rethinking the pattern of external policy referencing: media discourse over the ‘Asian tigers’ PISA success in Australia, Germany and South Korea. *Comparative Education*, 50(3), 302-321. DOI:10.1080/03050068.2013.860704.
- Thorndike, R. (1962). International Comparison of the Achievement of 13-year olds. In Foshay, A. *Educational Achievements of Thirteen Year Olds in Twelve Countries*. Hamburg: UNESCO Institute for Education.
- Trace, A.S. (1962). *What Ivan Knows and Johnny Doesn't*. New York: Random House.
- The Thomas B. Fordham Institute. (2015). About Fordham. Retrieved from www.edexcellence.net/about-us/fordham-mission.html.
- Tröhler, D. (2011). Concepts, Cultures and Comparisons. In Pereyra, M. et al (eds.) *PISA under examination: Changing knowledge, changing tests, and changing schools*. Boston, MA: Sense Publishers.
- Tyack, D. and Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, MA: Harvard University Press.
- Van Evera, S. (1997). *Guide to methods for students of political science*. Ithaca, NY: Cornell University Press.
- Vinovskis, M. (2009). *From A Nation at Risk to No Child Left Behind: National education goals and the creation of federal education policy*. New York: Teachers College Press.
- Waldow, F. (2009). Undeclared imports: silent borrowing in educational policy-making and research in Sweden. *Comparative Education*, 45 (4), 477-494. DOI: 10.1080/03050060903391628.
- Waldow, F. (2012). Standardisation and legitimacy: Two central concepts in research on educational borrowing and lending. In Steiner-Khamsi, G. and Waldow, F. (Eds). *World Yearbook of Education: Policy borrowing and lending in education*. New York: Routledge.
- Weiseltier, L. (2015, January 18). “Among the Disrupted.” *The New York Times Book Review*. p. BR1 and 14-15.
- Weiss, C. (1979). The many meanings of research utilization. *Public Administration Review*, 39(5), 426-431.

- Weyrauch, P. (2010, October 13-14). *Channelling (stet) expert-knowledge? How do political actors use expertise in parliamentary debates?* Paper presented at Making Parliaments Speak: method and issues in the analysis of debates within political assemblies, Paris, France.
- Whitty, G. (2012). Policy tourism and policy borrowing: A trans-Atlantic case study. In Steiner-Khamisi, G. and Waldow, F. (Eds). *Policy Borrowing and Lending in Education. World Yearbook of Education 2012*, 354-370. New York, NY: Routledge.
- Williams, J.H. and Engle, L.C. (2013). Testing to rank, testing to learn, testing to improve: an introduction and overview. *Research in Comparative and International Education*, 8(3), 214- 235. DOI: 10.2304/rcie.20138.3.214.
- Wiseman, A.W. (2013). Policy responses to PISA in comparative perspective. In Meyer, H-D. and Benavot, A. (Eds.) *PISA, power and policy: The emergence of global educational governance*. (303-322). Oxford, UK: Symposium Books.
- Yettick, H. (2009). *The research that reaches the public: Who produces the educational research mentioned in the news media?* Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. Retrieved from <http://epicpolicy.org/publication/research-that-reaches>.

Appendix A: Sampled Texts

Academic literature

- Archer, L, DeWitt, J, Osborne, J., Dillon, J., Willis, B., and Wong, B. (2012). Science aspirations, capital and family habitus: How families shape children's engagement and identification with science. *American Educational Research Journal*, 49, 881-908. DOI: 10.3102/0002831211433290.
- Baker, E. (2007). The end(s) of testing. *Educational Researcher*, 36, 309-317. DOI: 10.3102/001318907307970.
- Baumert, J., Kunter, M., Blum, W., Bruner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom and student progress. *American Educational Research Journal*, 47, 133-180. DOI: 10.3102/0002831209345157.
- Chiu, M. and Khoo, L. (2005). Effects of resources, inequality, and privilege bias upon achievement: Country, school and student level analyses. *American Educational Research Journal*, 42, 575-603. DOI: 10.3192/00028312042004575.
- Chudgar, A. and Luschei, T. (2009). National income, income inequality and the importance of schools: A hierarchical cross-national comparison. *American Educational Research Journal*, 46, 626-658. DOI: 10.3102/0002831209340043.
- Condrón, D. (2011). Egalitarianism and educational excellence: compatible goals for affluent societies? *Educational Researcher*, 40, 47-55. DOI: 10.3102/001318911401021.
- Dalton, B. (2012). Grade level and science achievement: US performance in cross-national perspective. *Comparative Education Review*, 56(1), 125-154. <http://www.jstor.org/stable/10.1086/660745>.
- Darling-Hammond, L. (2006). Securing the right to learn: Policy and practice for powerful teaching and learning. *Educational Researcher*, 35, 13-24. DOI: 10.3102/0013189X035007013.
- Dupriez, Duman and Vause. (2008). How do schools manage pupils heterogeneity? *Comparative Education Review*, 52(2), 245-273. <http://www.jstor.org/stable/10.1086/528764>.

- Gorard and Smith. (2004). An international comparison of equity in educational systems. *Comparative Education*, 40 (1), 15-28. DOI: 10.1080/0305006042000184863.
- Hanson, M. (2006). Transnational corporations as educational institutions of international development. *Comparative Education Review*, 50 (4), 625-650.
- Kramarski, B. and Mevarech, Z. (2003). Enhancing mathematical reasoning in the classroom: The effects of cooperative learning and metacognitive training. *American Educational Research Journal*, 40, 281-310. DOI: 10.3102/00028312040001281.
- Kamens, D. and McNeely, C. (2010). Globalization and the growth of international educational testing and international assessment. *Comparative Education Review*, 54(1), 5-25. DOI: 10.1086/648471.
- Leu, D., O'Byrne, I., Zawilinski, L., McVerry, J.G., Everett-Cacopardo, H. (2009). Comments on Greenhow, Robelia and Hughes: Expanding the new literacies conversation. *Educational Researcher*, 38, 264-269. DOI: 10.3102/0013189X09336676.
- Luke, A. (2011). Generalizing across borders: Policy and the limits of educational science. *Educational Researcher*, 40, 367-377. DOI: 10.3102/0013189x11424314.
- Luyten, H., Peschar, J. and Coe, R. (2008). Effects of schooling on reading performance, reading engagement and reading activities of 15-year-olds in England. *American Educational Research Journal*, 45, 319-342. DOI: 10.3102/0002831207313345.
- Ma, X. (2008). Within school gender gaps in reading, mathematics, and science literature. *Comparative Education Review*, 52 (3), 437-460.
- Marsh, H., Trautwein, U., Lüdtke, O., Baumert, J., Köller, O. (2007). The big-fish-little-pond effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal*, 44, 631-669. DOI: 10.3102/0002831207306728.
- McDaniel, A. (2010). Cross national gender gaps in educational expectations. *Comparative Education Review*, 54(1), 27-50. DOI: 10.1086/648060.
- Möller, J., Tetelsdorf, J., Köller, O., Marsh, H. (2011). The reciprocal internal/external frame of reference model: An integration of models of relations between

- academic achievement and self-concept. *American Educational Research Journal*, 48, 1315-1345. DOI: 10.3102/0002831211419649.
- Park, H. (2008). The varied educational effects of parent-child communications. *Comparative Education Review*, 52(2), 219-243.
- Pizmony-Levy, O. (2011). Bridging the global and the local in understanding curricula scripts. *Comparative Education Review*, 55(4), 600-633. <http://www.jstor.org/stable/10.1086/661632>.
- Robertson, S. (2012). Placing teachers in global governance agendas. *Comparative Education Review*, 56(4), 584-607. <http://www.jstor.org/stable/10.1086/667414>.
- Robinson, J.P. and Lubienski, S.T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, 48, 268-302. DOI: 10.3102/0002831210372249.
- Rupp, A. and Lesaux, N. (2006). Meeting Expectations? An empirical investigation of a standards-based assessment of reading comprehension. *Educational Evaluation and Policy Analysis*, 28, 315-333. DOI: 10.3102/01623737028004315.
- Rutkowski, L., Gonzalez, E., Joncas, M., and von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39, 142-151. DOI: 10.3102/0013189X10363170.
- Schmidt, W. and Hoang, R. (2012). Curricular coherence and the Common Core State Standards for mathematics. *Educational Researcher*, 41, 294-308. DOI: 10.3102/0013189X12464517.
- Seaton, M., Marsh, H., and Craven, R. (2010). Big-fish-little-pond effect: Generalizability and moderation – two sides of the same coin. *American Educational Research Journal*, 47, 390-433. DOI: 10.3102/0002831209350493.
- Shenkar, M. and Shenkar, O. (2011). Labor conflict on the national stage: Metaphoric lenses in Israel's teachers strike. *Comparative Education Review*, 55(2), 210-230. <http://www.jstor.org/stable/10.1086/657973>.
- Simola, H. (2005). The Finnish miracle of PISA: Historical and sociological remarks on teaching and teacher education. *Comparative Education*, 41(4), 455-470. DOI: 10.1080/03050060500317810.
- Steiner-Khamsi, G. (2010). The politics and economics of comparison. *Comparative Education Review*, 54(3), 323-342. <http://www.jstor.org/stable/10.1086/653047>.

- Stone, J.R., III, Alfeld, C. and Pearson, D. (2008). Rigor and relevance: Enhancing high school students' math skills through career and technical education. *American Educational Research Journal*, 45, 767-795. DOI: 10.3102/0002831208317460.
- Takayama, K. (2008). The politics of international league tables: PISA in Japan's achievement crisis debate. *Comparative Education*, 44(4), 387-407. DOI: 10.1080/0305006802481413.
- Takayama, K. (2009). Politics of externalization in reflexive times: Reinventing Japanese education reform discourses through "Finnish PISA success." *Comparative Education Review*, 54(1), 51- 75. DOI: 10.1086/644838.
- Vaughn, S., Klingner, J.K., Swanson, E.A., Boardman, A.G., Roberts, G., Mohammed, S.S., et al. (2011). Efficacy of collaborative strategic reading with middle school students. *American Educational Research Journal*, 48, 938-964. DOI: 10.3102/0002831211410305.
- Xu and Hampden-Thompson. (2012). Cultural reproduction, cultural mobility, cultural resources or trivial effect: A comparative approach to cultural capital and educational performance. *Comparative Education Review*, 56, 98-124. <http://www.jstor.org/stable/10.1086/661289>.
- Xu, J. (2008). Sibship size and educational achievement: The role of welfare regimes cross-nationally. *Comparative Education Review*, 52(3), 413-436. <http://www.jstor.org/stable/10.1086/588761>.

Think tank literature

- A certain Finnish je ne sais quoi. (2008, March 5). Retrieved from <http://edexcellence.net/commentary/education-gadfly-weekly/2008/march-6/a-certain-finnish-je-ne-sais-quoi.html>.
- Bungalawala, Z. (2011). Young, educated, and dependent on the public sector. Doha: Brookings Doha Center. Retrieved from <http://www.brookings.edu/research/papers/2011/12/15-qatar-diversify-employment>.
- Cardenas, M. (2011, February 23). The 2010s: Latin America's fleeting opportunity. Retrieved from <http://www.brookings.edu/research/opinions/2011/02/23-fleeting-opportunity-cardenas>.
- Carmichael, S., Wilson, S.W., Finn, C., Winkler, A., and Palmieri, S. (2009). Stars by which to Navigate? Scanning National and International Education Standards in

2009. Washington, DC: The Thomas B. Fordham Institute. Retrieved from <http://edexcellence.net/publications/stars-by-which-to-navigate.html>.
- Fairchild, D. (2012, December 12). A reverse Sputnik moment? Retrieved from <http://edexcellence.net/commentary/education-gadfly-weekly/2012/december-13/a-reverse-sputnik-moment.html>.
- Finn, C.E. Jr., (2010, December 8). A Sputnik moment for US education. Retrieved from <http://edexcellence.net/commentary/education-gadfly-weekly/2010/december-9/a-sputnik-moment-for-u-s-education.html>.
- Greenstone, M. and Looney, A. (2011, September 22). Are we shortchanging our future? The economic imperative of attracting great teachers. Retrieved from http://www.hamiltonproject.org/files/downloads_and_links/092211_teachers_wages.pdf.
- Haskins, R., Murnane, R., Sawhill, I., and Snow, C. (2012, October 12). Can Academic Standards Boost Literacy and Close the Achievement Gap? Washington, DC: Center on Children and Families, Brookings Institute. Retrieved from <http://www.brookings.edu/research/papers/2012/10/02-boost-literacy-haskins-sawhill>.
- Joch, N. (2010, December 9). The Chinese delegation. Retrieved from <http://edexcellence.net/commentary/education-gadfly-daily/flypaper/2010/the-chinese-delegation-to-e2-80-a6-ohio.html>.
- Kraetsch, M. and Constant, S. (2010, June). Taking stock of the youth challenge in the Middle East: New data and new questions. Retrieved from <http://www.brookings.edu/research/articles/2010/06/middle-east-youth>.
- Loveless, T. (2009). The 2008 Brown Center Report on American Education: How Well Are American Students Learning? Washington, DC: Brookings Institute. Retrieved from <http://www.brookings.edu/research/reports/2009/02/25-education-loveless>.
- Loveless, T. (2011). The 2010 Brown Center Report on American Education: How Well Are American Students Learning? Washington, DC: Brookings Institute. Retrieved from <http://www.brookings.edu/research/reports/2011/02/07-education-loveless>.
- Negroponte, D. (2012, November 30). Mexico's Enrique Pena Nieto Confronts the Challenges of Federalism and Fiscal Reform and Education. Retrieved from <http://www.brookings.edu/blogs/up-front/posts/2012/11/30-mexico-pena-nieto-negroponte>.

- Northern, A. (2009, February 26). PISA smackdown. Retrieved from <http://edexcellence.net/commentary/education-gadfly-daily/flypaper/2009/pisa-smackdown.html>.
- Petrilli, M. and Scull, J. (2011). American Achievement in International Perspective. Washington, DC: The Thomas B. Fordham Institute. Retrieved from <http://edexcellence.net/publications/american-achievement-in.html>.
- Porter-Magee, K. (2012, December 27). Real lessons from Finland: Hard choices rigorously implemented. Retrieved from <http://edexcellence.net/commentary/education-gadfly-daily/flypaper/2012/real-lessons-from-finland-hard-choices-rigorously-implemented.html>.
- Puryear, J. and Goodspeed, T. (2012). Education in the Americas: What the summit missed. In *The Road to Hemispheric Cooperation: Beyond the Cartagena Summit of the Americas*. Washington, DC: Brookings Institute. Retrieved from <http://www.brookings.edu/research/reports/2012/07/07-summit-of-the-americas>.
- van Fleet, J. and Zinny, G. (2012). Corporate Social Investments in Education in Latin America and the Caribbean: Mapping the magnitude of multinationals' private dollars for education. Washington, DC: The Brookings Institute. Retrieved from <http://www.brookings.edu/research/papers/2012/08/investments-latin-america-van-fleet>.
- Wagner, D., Murphy, K., and DeKorne, H. (2012). Learning First: A Research Agenda for Improving Learning in Low-Income Countries. Center for Universal Education Working Paper. Retrieved from <http://www.brookings.edu/research/papers/2012/12/learning-first-wagner-murphy-de-korne>.
- Winkler, A., Ballard, A. and Palmieri, S. (2009). Education Olympics 2008: The Games in Review. Washington, DC: The Thomas B. Fordham Institute. Retrieved from <http://edexcellence.net/publications/education-olympics-2008-the.html>.

Media

The New York Times

- Anderson, J. (2011, December 12). From Finland, and intriguing school-reform model. *The New York Times*. p. A33.

- AP, (2007, December 5). Other Countries students' surpass U.S.'s on Tests. *The New York Times*. Retrieved from <http://www.nytimes.com/2007/12/05/education/05scores.html>.
- Arenson, K. (2004, December 15). Math and science tests find 4th and 8th graders still lag many peers. *The New York Times*. p. A29.
- Dillon, S. (2010, December 7). Top test scores from Shanghai stun educators. *The New York Times*. p. A22.
- Dillon, S. (2011, March 16). US must raise status of its teachers. *The New York Times*. p. A22.
- Friedman, T. (2009, April 21). Swimming without a suit. *The New York Times*. p. A27.
- Friedman, T. (2010, December 8). Skill digging. *The New York Times*. Retrieved from <http://www.nytimes.com/2010/12/08/opinion/08friedman.html>.
- Friedman, T. (2011, November 20). How about better parents? *The New York Times*. p. SR11.
- Friedman, T. (2012, August 8). Average is over, Part II. *The New York Times*. p. A19.
- Friedman, T. (2012, September 9). New rules. *The New York Times*. p. SR13.
- Friedman, T. (2012, 10 March). Pass the books. Hold the oil. *The New York Times*. p. SR1.
- Guttenplan, D.D. (2010, December 8). Western nations react to poor education results. *The New York Times*. Retrieved from <http://www.nytimes.com/2010/12/09/world/europe/09education.html>.
- Guttenplan, D.D. (2012, November 12). Comparing the success of nations. *The New York Times*. p. A21.
- Johnston, D. (2004, March 19). An economic and social imperative: Education needs to adapt to a changing world. *The New York Times*. Retrieved from http://www.nytimes.com/2004/03/19/opinion/19iht-edjohnston_ed3_.html.
- Kristof, N. (2011, January 15). China's Education System. *The New York Times*. Retrieved from <http://kristof.blogs.nytimes.com/2011/01/15/chinas-education-system/>.
- Leonhardt, D. (2011, June 8). The German example. *The New York Times*. p. B1.

- Pfeiffer, A. (2009, 29 April). O.E.C.D. Study maps shifting ideals of excellence. *New York Times*. Retrieved from <http://www.nytimes.com/2009/04/29/world/europe/29iht-riedpisa.html>.
- Rampell, C. (2009, April 22). Overpaying for educational underachievement. *The New York Times*. Online at <http://economix.blogs.nytimes.com/2009/04/22/overpaying-for-educational-underachievement>.
- Rampell, C. (2010, December 8). Education spending and those international test scores. *The New York Times*. Retrieved from <http://economix.blogs.nytimes.com/2010/12/08/education-spending-and-those-international-test-scores/>.
- Schmid, J. (2003, February 7). Even German students aren't doing well. *The New York Times*. Retrieved from http://www.nytimes.com/2003/02/07/news/07iht-schools_ed3_.html.
- Williams, T. (2011, August 1). A geography report. *The New York Times*. p. A14.
- The Wall Street Journal*
- Etter, L. (2010, December 7). U.S. News: American teens trail global peers in math scores – but US students show progress in science. *The Wall Street Journal*. p. A5.
- Finn, Chester, E., Jr. (2010, December 8). A Sputnik moment for US education. *The Wall Street Journal*. p. A21.
- Kronholz, J. (2004, June 7). Economic time bomb: US teens are among the worst at math. *The Wall Street Journal*. p. B1.
- Levine, A. (2012, November 15). The suburban education gap. *The Wall Street Journal*. p. A19.
- No passing grade. (2004, December 7). *The Wall Street Journal*. p. A10.
- Schultz, G. and Hanushek, E. (2012, May 1). Education is key to a healthy economy. *The Wall Street Journal*. p. A15.
- Tomsho, R. (2009, April 22). Study tallies education gaps effect. *The Wall Street Journal*. Retrieved from <http://www.wsj.com/articles/SB124040633530943487>.

Banchero, S. (2011, November 12). Brightest stall, low achievers gain. *The Wall Street Journal*. Retrieved from <http://www.wsj.com/articles/SB10001424052970203537304577032091650691280>.

The Washington Post

Ayers, R. (2010, September 27). What Superman got wrong, point-by-point. *The Washington Post*. (blog post) Retrieved from <http://voices.washingtonpost.com/answer-sheet/guest-bloggers/what-superman-got-wrong-point.html>.

Bessie, A. (2011, May 9). Our cutthroat curriculum. *The Washington Post*. (blog post) Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/our-cutthroat-curriculum/2011/05/05/AFuiFP3F_blog.html.

Chandler, M.A. (2008, July 13). Ads hope to inject US school challenges in to White House race. *The Washington Post*. p. C7.

Cuban, L. (2011, August 27). International rankings reformers like to ignore. *The Washington Post*. (blog post) Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/international-rankings-reformers-like-to-ignore/2011/08/27/gIQA6Gc6iJ_blog.html.

Darling-Hammond, L. (2012, April 27). Education and the income gap. *The Washington Post*. (blog post) Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/education-and-the-income-gap-darling-hammond/2012/04/26/gIQAhn0LkT_blog.html.

Dobbs, M. (2004, December 15). Good news, bad news about US education. *The Washington Post*. p. A02.

Dobbs, M. (2004, December 7). In a global test of math skills, US students behind the curve. *The Washington Post*. p. A01.

Eddy, M. (2003, August 24). German schools draw a barrage of criticism. *The Washington Post*. p. A21.

Glod, M. (2007, December 5). US teens trail peers around the world on math-science test. *The Washington Post*. p. A07.

Harris, G. (2012, December 10). Michelle Rhee wins over a critic. *The Washington Post*. (blog post) Retrieved from <http://www.washingtonpost.com/blogs/answer-sheet/wp/2012/12/08/michelle-rhee-wins-over-a-critic/>.

- Hatch, T. (2010, December 31). What Norway (not Finland) tells us about schools. *The Washington Post*. (blog post) Retrieved from <http://voices.washingtonpost.com/answer-sheet/guest-bloggers/what-norway-not-finland-tells.html>.
- Huffman, K. (2010, January 2). How to get top grades in education. *The Washington Post*. (blog post) Retrieved from <http://www.washingtonpost.com/wp-dyn/content/article/2010/01/01/AR2010010101416.html>.
- Kahlenberg, R. (2011, February 28). Debating Michelle Rhee. *The Washington Post*. (blog post) Retrieved from <http://voices.washingtonpost.com/answer-sheet/guest-bloggers/debating-michelle-rhee.html>.
- Kaiser, R. (2005, 24 May). Focus on schools helps Finns build a showcase nation. *The Washington Post*. p. A12.
- Kaiser, R. (2005, 7 August). In Finland's footsteps; If we're so rich and smart, why aren't we more like them? *The Washington Post*. p. B01.
- Lucido, H. (2011, May 21). Ten things teachers need to reclaim their profession. *The Washington Post*. (blog post) Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/ten-things-teachers-need-to-reclaim-their-profession/2011/05/15/AFeo3V8G_blog.html?wprss=answer-sheet.
- Matthews, J. (2009, 19 October). Tests don't always offer right answers. *The Washington Post*. B03.
- Matthews, J. (2011, August 17). Let's pick the world's best education nation. *The Washington Post*. www.washingtonpost.com.
- Matthews, J. (2011, February 11). The myth of declining US schools: They've long been mediocre. *The Washington Post*. (blog post) Retrieved from http://voices.washingtonpost.com/class-struggle/2011/02/myth_of_declining_us_schools.html.
- Page, B. (2012, June 4). The dangers of not teaching personal finance to kids. *The Washington Post*. (blog post) Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/the-dangers-of-not-teaching-personal-finance-to-kids/2012/06/03/gJQAZchiBV_blog.html.
- Porter, A. (2012, August 27). Back to school: Rethinking America's global education rankings. *The Washington Post*. Retrieved from

- http://www.washingtonpost.com/blogs/therootdc/post/back-to-school-rethinking-america-s-global-education-rankings/2012/08/27/af11b510-f06e-11e1-ba17-c7bb037a1d5b_blog.html.
- Ravitch, D. (2011, October 13). Why Finland's schools are doing great (by doing what we don't). *The Washington Post*. (blog post) Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/ravitch-why-finlands-schools-are-great-by-doing-what-we-dont/2011/10/12/gIQAmTyLgL_blog.html.
- Salhberg, P. (2012, April 17). What the US can't learn from Finland about education reform. *The Washington Post*. (blog post) Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/what-the-us-cant-learn-from-finland-about-ed-reform/2012/04/16/gIQAGIvVMT_blog.html.
- Samuelson, R. (2011, January 10). An achievement gap that won't be fixed in schools. *The Washington Post*. p. A17.
- Sener, J. (2011, April 10). I'm better than Jordan. The test results say so. *The Washington Post*. (blog post) Retrieved from http://www.washingtonpost.com/opinions/standardized-tests-prove-im-better-than-michael-jordan/2011/03/29/AF4sdL4C_story.html.
- Slade, S. (2010, November 2). What other countries are really doing in education. *The Washington Post*. (blog post) Retrieved from <http://voices.washingtonpost.com/answer-sheet/guest-bloggers/what-other-countries-are-reall.html>.
- Strauss, V. (2010, December 6). Do international test comparisons make sense? *The Washington Post*. (blog post) Retrieved from <http://voices.washingtonpost.com/answer-sheet/standardized-tests/so-what-if-the-us-is-not-no-1.html>.
- Strauss, V. (2010, December 7). Hysteria over PISA misses the point. *The Washington Post*. (blog post) Retrieved from <http://voices.washingtonpost.com/answer-sheet/standardized-tests/hysteria-over-pisa.html>.
- Strauss, V. (2010, December 7). Sample PISA questions given to 15-year olds. *The Washington Post*. (blog post) Retrieved from <http://voices.washingtonpost.com/answer-sheet/standardized-tests/sample-pisa-questions-given-to.html>.
- Strauss, V. (2011, February 16). Mobile technology and school reform: what's the connection. *The Washington Post*. (blog post) Retrieved from

<http://voices.washingtonpost.com/answer-sheet/technology/mobile-technology-and-school-r.html>.

Strauss, V. (2012, March 12). The Joel Klein-Condi Rice ed report: What it will and won't say. *The Washington Post*. (blog post). Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/the-joel-klein-condi-rice-ed-report-what-it-will-and-wont-say/2012/03/03/gIQA21Pp7R_blog.html.

Strauss, V. (2012, March 21). Best part of 'schools-threaten-national-security' report: The dissents. *The Washington Post*. (blog post). Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/best-part-of-schools-threaten-national-security-report-the-dissents/2012/03/20/gIQAe0yaQS_blog.html.

Thomas, P. (2011, May 14). Bill Maher's "Real Time" education. *The Washington Post*. (blog post) Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/bill-mahers-real-time-education-debate-failure/2011/05/14/AF0EvU3G_blog.html.

Thomas, P. (2011, April 5). A case against standards. *The Washington Post*. (blog post). Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/post/2011/04/04/AFgXSFgC_blog.html.

Welner, K. (2010, December 11). If you find yourself in a hole, stop digging. *The Washington Post*. (blog post) Retrieved from <http://voices.washingtonpost.com/answer-sheet/guest-bloggers/if-you-find-yourself-in-a-hole.html>.

Willingham, D. (2010, December 13). What causes performance decline across grads? *The Washington Post*. (blog post) Retrieved from <http://voices.washingtonpost.com/answer-sheet/daniel-willingham/willingham-whata-causes-declin.html>.

Yang, J.L. (2012, November 18). The signs of America's decline. *The Washington Post*. p. B01.

Zhao, Y. (2012, December 12). Yes, numbers (on international tests) can lie. *The Washington Post*. (blog post) Retrieved from <http://www.washingtonpost.com/blogs/answer-sheet/wp/2012/12/12/yes-numbers-on-international-tests-can-lie/>.

Zhao, Y. and Ravitch, D. (2010, December 14). How Shanghai topped PISA rankings and why it's not big news in China. *The Washington Post*. (Blog post) Retrieved

from <http://voices.washingtonpost.com/answer-sheet/standardized-tests/how-shanghai-topped-pisa-ranki.html>.

USA Today

Chua, A. (2011, May 11). How to reshape US education. *USA Today*. p. A9.

Henry, T. (2001, December 4). Getting a read on US teens. *USA Today*. p. D08.

Lynch, D. (2008, November 19). Learning from Korean schools: US could use some tips on preserving competitiveness. *USA Today*. p. B1.

Neill, M. and Guisbond, L. (2004, December 15). Rethink standardized tests. *USA Today*. A19.

Toppo, G. (2004, December 7). US teens have weak practical math skill. *USA Today*. p. D7.

van Roekel, D. (2010, December 15). Elevate the teaching profession. *USA Today*. p. A8.