Barriers to Implementation of New Programs and Pedagogies in K-12 STEM Education: A Systems Perspective

by

Mackenzie Douglas Hird

B.S., Physics
University of Texas at Dallas, 2011

Submitted to the Engineering Systems Division
in Partial Fulfillment of the Requirements for the Degree of Master of Science in Technology and Policy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

© 2013 Massachusetts Institute of Technology
All Rights Reserved

Signature of Author

Technology and Policy Program
Engineering Systems Division
May 10, 2013

Certified by

Richard Larson
Mitsui Professor of Engineering Systems
Thesis Supervisor

Accepted by

Dava J. Newman
Professor of Aeronautics and Astronautics and Engineering Systems Director, Technology and Policy Program
Barriers to Implementation of New Programs and Pedagogies in K-12 STEM Education: A Systems Perspective

by

Mackenzie Douglas Hird

Submitted to the Engineering Systems Division on May 10, 2013
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Technology and Policy

Abstract

The continued usage of poor pedagogies in K-12 classrooms, despite large pressures for teachers to change their practice, points towards systematic barriers to change. In the last few decades, there has been a national focus to improve Science, Technology, Engineering and Math (STEM) Education. Driven by their concern for developing their future workforce, science and technology companies have invested billions of dollars in improving student outcomes. Further, the federal and state governments have responded by adopting new policies meant to improve student performance. Promising new pedagogies, such as Project Based Learning or the Flipped Classroom, have been developed alongside new technologies to complement them. Yet despite this support, pedagogical practice has not drastically changed and students are primarily taught through lectures and homework sets.

This thesis argues that teachers do not adopt new pedagogies because they are under short-term pressure to improve test scores, often face an uphill battle against their school culture and/or do not have deep enough pedagogical or content expertise. A causal model of pedagogical implementation barriers is developed using the results of in-depth surveys and interviews of administrators, principals and teachers. Within this model, critical points of leverage are identified that can interrupt the negative feedback loops creating pedagogical lock-in, and three case studies of international attempts at pedagogical reform are presented to illustrate effective strategies to utilize these leverage points. General policy recommendations are then developed that will remove the current system of pressures and incentives for teachers to use rote memorization and incentivize use of more effective pedagogies.

Thesis Supervisor: Richard Larson
Title: Mitsui Professor of Engineering Systems
Acknowledgments

I would first like to express my gratitude to Professor Richard Larson, who has been an incredible mentor over the last two years. His advice and support in my development as a policy analyst and more effective communicator, as well as his ability to reduce complex systems down to their most critical components, have been invaluable during this work. His passion for improving education is simply contagious and I hope that passion shines through in this thesis.

Some of the materials related to Lesson Study and cross-cultural comparisons represent joint work with Dr. Yuko Okubo of Fujitsu Laboratories of America. These contributions are seen in their entirety in the following separate presentations: Okubo, Hird, Uchino and Larson, “An Examination of Learning Standards in a Comparative Perspective: U.S., Japan, Singapore, and Finland,” Comparative and International Education Society Annual Conference, 2013 and Hird, Okubo, and Uchino, “Integration of Technology Enabled Education in Learning: a comparative study of the influence of learning standards in the U.S., Japan, Singapore, and Finland,” Learning International Networks Consortium, 2013. I thank Yuko sincerely for being my research colleague and for allowing me to include such joint work in this thesis.

My colleagues in the Education as a Complex System group have been a source of constant support. I can always count on them for an engaging discussion or helpful suggestion. I would especially like to thank Navid Ghaffarzadegan, whose advice in refining my approach to developing the model was instrumental in this work.

The faculty and staff of both the Technology and Policy Program and Engineering Systems Division always had their doors open and were always ready for a laugh. My relationships with my fellow students have been both rewarding and inspiring. Who knew you all did such great research and were fun people?

Much of my education in the last few years has happened outside of a university, helping me to see the broader context of my work. I am especially thankful to both Heidi Schweingruber at the National Research Council and Toby Smith at the Association of American Universities for their valuable mentorship during my forays into education policy in Washington, DC. Many great experiences (and controlled explosions in front of a classroom of students) with Ben Freitag and the rest of the EYES staff were what got me interested in Science Education in the first place. I would also like to thank my McDermott and UT Dallas Family, along with some incredible teachers at Luther College High School, who have helped me get to where I am today.

Finally I would like to thank my family and friends, who were always a source of both laughter and support. My parents and brother have been always been a source of encouragement, helping to motivate me every day and keep me on track. I would especially like to thank Dina for always being right beside me every step of the way over the last two years, and especially throughout the last few weeks.

The research leading to this paper was supported by Fujitsu Laboratories of America under contract to the Massachusetts Institute of Technology, called, “Towards Intelligent Societies: What Motivates Students to Study Science and Math? How Can We Provide Flexible Learning Pathways?” The opinions expressed in this thesis represent those of the author and do not necessarily represent the opinions of Fujitsu Laboratories of America or the Massachusetts Institute of Technology.
# Table of Contents

Acknowledgments ................................................................................................................................................ 5

Table of Contents ................................................................................................................................................ 7

List of Figures .................................................................................................................................................. 9

1. Introduction ................................................................................................................................................... 11
   1.1 Research Question .................................................................................................................................. 12

2. Methods ......................................................................................................................................................... 17
   2.1 Research Design ...................................................................................................................................... 17
   2.2 Case Selection ........................................................................................................................................ 19
   2.3 Data Collection ....................................................................................................................................... 21
   2.4 Measurement of Key Variables ............................................................................................................... 22
   2.5 Data Analysis ......................................................................................................................................... 24

3. Literature Review .......................................................................................................................................... 25
   3.1 Improved Pedagogies Necessary for Higher Student Achievement ....................................................... 25
   3.2 Mismatch between Test Scores and Student Knowledge ......................................................................... 28
   3.3 School Culture Impedes New Pedagogies .............................................................................................. 30
   3.4 Teacher Workload Impedes Use of New Pedagogies ............................................................................ 32
   3.5 Systems Level Miscommunications ...................................................................................................... 34
   3.6 Glossary of Key Terms .......................................................................................................................... 34

4. Results and Model Development ................................................................................................................ 37
   4.1 Contradictory Survey Results .................................................................................................................. 37
   4.2 Administrator Rationale behind Assessment-Based Accountability ....................................................... 39
   4.3 Shortcut to Improve Test Scores ........................................................................................................... 43
   4.4 Culture Reinforcing Teaching to the Test ............................................................................................... 46
   4.5 The Effect of Content Knowledge on Pedagogy .................................................................................... 49
   4.6 The Effect of Pedagogical Knowledge .................................................................................................... 51
   4.7 Full Causal Model of Factors that Effect Pedagogy ............................................................................... 53

5. Leverage Points and International Comparisons ......................................................................................... 55
   5.1 Leverage Points ....................................................................................................................................... 55
      5.1.1 Assessment Design and Objectivity of Measures ............................................................................. 55
      5.1.2 Teacher Pedagogical Knowledge .................................................................................................... 59
      5.1.3 Teacher Content Knowledge ........................................................................................................ 63
      5.1.4 Culture and Competition ................................................................................................................. 64
   5.2 International Use of Leverage Points to Affect Change ........................................................................ 65
      5.2.1 The Japanese Experience ............................................................................................................... 66
      5.2.2 The Singaporean Experience .......................................................................................................... 69
      5.2.3 The Finnish Experience .................................................................................................................. 72

6. Policy Recommendations ............................................................................................................................... 77
   6.1 Impact of the Next Generation Science Standards ................................................................................ 77
   6.2 Generalized Policy Recommendations .................................................................................................. 79
      6.2.1 Remove Standardized Tests and Develop a Legitimating System for Subjective Tests .................... 80
      6.2.2 Implement Lesson Study as Improved Pedagogical Training ......................................................... 80
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2.3 Improve Pedagogical Practice at Universities to Improve Future Teacher Knowledge</td>
<td>81</td>
</tr>
<tr>
<td>6.2.4 Change the Culture in Schools through Pre-service Clinical Experiences</td>
<td>81</td>
</tr>
<tr>
<td>6.3 Future Work</td>
<td>82</td>
</tr>
<tr>
<td>6.4 Conclusions</td>
<td>83</td>
</tr>
<tr>
<td>Bibliography</td>
<td>85</td>
</tr>
<tr>
<td>Appendix A: Survey Instrument</td>
<td>99</td>
</tr>
<tr>
<td>Appendix B: Administrators and Principals Interview Protocol</td>
<td>101</td>
</tr>
<tr>
<td>Appendix C: Teacher Interview Protocol</td>
<td>103</td>
</tr>
<tr>
<td>Appendix D: Survey Results</td>
<td>105</td>
</tr>
<tr>
<td>Appendix E: Time Spent Teaching across OECD Countries</td>
<td>109</td>
</tr>
</tbody>
</table>
**List of Figures**

Figure 1: Time spent Teaching Across OECD Countries ................................................. 33
Figure 2: Rationale for Assessment Based Accountability ................................................. 42
Figure 3: Society Level Rationale for Controlling Teacher Performance ......................... 43
Figure 4: Shortcut to improve test scores ........................................................................ 45
Figure 5: Long Term Effects of Rote Memorization ......................................................... 46
Figure 6: Reinforcing teaching to the test within the full model ....................................... 49
Figure 7: The Effect of Content Knowledge ..................................................................... 51
Figure 8: Effect of Previous Training on Teaching Ability ............................................... 53
Figure 9: Japanese Attempts with Reforms to use new Pedagogies ................................. 69
Figure 10: Singaporean Attempts with Reforms to use new Pedagogies ......................... 72
Figure 11: Finnish Attempts with Reforms to use new Pedagogies ................................. 75
Figure 12: Unchanged Leverage Points under NGSS that will impede changes in pedagogy .... 79
1. Introduction

Scientific research and development has been a key driver of the American economy since the end of World War II. From 1950 to 1998, technical progress accounted for 58% of the growth in American GDP.¹ This progress was supported by a strong scientific workforce trained by an education system consisting of world-class elementary, secondary, and higher education institutions.² However, there are troubling signs that this scientific excellence might be slipping away as the US falls behind other countries in primary and secondary STEM education. On the 2009 PISA exam, internationally recognized as a high quality assessment of student achievement across the 34 OECD countries, the United States was ranked 17th in science and 25th in mathematics performance.³ Without a strong system of science education developing a scientifically literate workforce, many policymakers are concerned that science and technology will not drive the United States economy for much longer.

In response to these concerns of stagnating performance of schools, federal, state and local governments have made significant policy changes designed to radically improve the Science, Technology, Engineering and Math (STEM) education system. The federal government passed the No Child Left Behind Act in an effort to hold schools accountable for their behavior, and more recently poured billions of dollars into schools and districts around the country through its Race To The Top competition to incentivize the adoption of new policies. States have launched hundreds of individual pilot programs to experiment with new methods of teaching. Science, technology and engineering companies, recognizing their own self-interest in the continued development of a highly-skilled workforce, have spent millions of dollars to develop new ways of teaching and new materials to be used in classrooms. Yet none of these efforts have created the radical changes in STEM learning that were envisioned. Test scores haven’t improved. The number of students entering STEM careers hasn’t increased.

² President’s Council of Advisors on Science and Technology, Prepare and Inspire: K-12 Education in Science, Technology, Engineering and Math (STEM) for America’s Future, (The White House Office of Science and Technology Policy: Washington, DC, 2010).
Policymakers and education researchers have identified many potential reasons for such results, including lack of funding, poor textbooks created by publishers, bloated administrative structure or teacher union resistance to change. However, there is a growing body of evidence pointing to the continued usage of poor pedagogies as the primary reason behind stagnating test scores. More recently, states have developed the Next Generation Science Standards. These standards strive to promote the adoption of pedagogies for an exploratory scientific education rather than memorization of facts and figures. However, it is difficult to understand how the Next Generation Science Standards might change the adoption of new pedagogies without a deep understanding of how the current system incentivizes and reinforces the use of particular pedagogies. This thesis will examine the processes responsible for the continued usage of lecture-based teaching, despite the large base of evidence supporting the use of new pedagogies.

1.1 Research Question

The difficulty identifying the problems with implementing improved pedagogies is that there are actually numerous answers to this question. As we will see in the literature review, each study typically identifies one major factor impeding the implementation of a new pedagogy. However, the education system is exactly that: a system. Each structural element depends on one another, and changing one element changes the others as well. For this reason, the National Academy of Education has concluded that there must be a coherent approach to education reform across all levels of the education system. The same is true of any policy change that attempts to move teachers towards adopting new pedagogies. Without understanding the underlying dynamics of the education system, these policies will not be effective as teachers continually run into barriers.

This problem, therefore, is especially ripe for the application of a Systems Dynamics model capable of modeling such complex interactions. However, the first and most important step of designing a Systems Dynamics model is to understand the underlying causal relationships and feedback loops. This work focuses on the development of such causal relationships. By

---

examining the actions of each of the critical occupations in the education system, (that is teachers, principals and administrators/bureaucrats), this thesis answers the following question:

*How does the structure and interactions of the complex education system cause problems for the implementation of new pedagogies or technologies in STEM education?*

The focus of the work is to develop a robust model to describe the barriers to change caused by the structure of the K-12 STEM education system. While this project does not claim to be comprehensive enough to explain every interaction of the extremely complex system, it has produced valuable insight for overcoming the resistance to change.

This is not the first application of such a Systems Dynamics model to STEM Education. The Business Higher Education Forum, with the help of Raytheon, has built their U.S. STEM Education Model over the course of many years of effort. This model is extremely comprehensive in scope and attempts to understand the dynamics of how and why students pursue a career in STEM. This has been an exciting first step in using these modeling techniques for a novel purpose, but it has not yet generated the kind of insight that was initially expected. Furthermore, critics argue that the recommendations developed by this model are so broad that they add little to the discourse other than reinforce existing beliefs, such as “increasing student interest will increase the number of STEM graduates.” Smaller, more focused models have been successful in modeling certain sections of the education system. For instance, Sturtevant’s work on modeling the teacher workforce has been successful in demonstrating the unintended consequences of high STEM worker pay drawing well-qualified professionals from teaching. With such smaller models, we can generate an intimate understanding of the situation that can lead to valuable insight.

Also guiding the design of this model is the large gap between most policy recommendations and the “on-the-ground” needs of schools. In STEM Education there are typically two levels of research undertaken: broad nationwide recommendations, and the assessments of specific programs in specific schools. Two of the most influential reports in STEM education, *Prepare and Inspire* and *The Opportunity Equation*, provide only broad

---

nation-wide recommendations such as “use more inquiry base learning methods”\textsuperscript{9,10} and “create 1,000 new STEM-focused schools over the next decade”\textsuperscript{11,12} without much on-the-ground guidance for educators. Other relevant literature, for instance on using inquiry-based learning, deals with the use of specific proprietary tools, typically technology-based, for a specific type of learning\textsuperscript{13} or in a specific school.\textsuperscript{14} While both broad and specific types of literature are essential, there is no work that combines both and provides broad guidance for administrators, principals and teachers on how to best implement such programs across all school contexts. This thesis strives to consider those factors and build a model that is generalizable to provide a much needed decision making tool for those designing policy or technologies to improve education. This type of work, where broad generalizability is the goal, is often difficult in education because of the multitude of confounding factors that influence what goes on in each school, such as socioeconomic status, prior achievement of students, quality of teacher, etc.

We approached this work with a set of more specific hypotheses that we believe may be true of the causal system behind decisions in STEM Education:

- Teachers do not choose to use outdated teaching methodologies, such as teaching by rote memorization, because they believe it increases student learning; rather, they use these methods because they face large consequences for not improving test scores
- There is little pursuit of alternate pedagogies by teachers because there is no demand from students, parents or administrators. Parents and students are focused on what they need to do to obtain good grades, graduate, or pass the class; not for the love of learning itself

\textsuperscript{9} President’s Council of Advisors on Science and Technology, \textit{Prepare and Inspire: K-12 Education in Science, Technology, Engineering and Math (STEM) for America’s Future.}
\textsuperscript{11} President’s Council of Advisors on Science and Technology, \textit{Prepare and Inspire: K-12 Education in Science, Technology, Engineering and Math (STEM) for America’s Future.}
\textsuperscript{12} National Research Council, \textit{Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future.}
• Inadequate teacher preparation and support lead those few teachers who are still teaching after their first few years to develop informal coping habits (pedagogical lock-in, group lesson planning, online resources etc.) that are difficult to systemically change

• The broad STEM curriculum makes it difficult for STEM teachers to be experts in everything they are expected to teach, and as such teachers are hesitant to conduct inquiry-based lessons that will place them outside of their comfort zone

The next section, Chapter 2, will detail the method used to survey and interview educators. Chapter 3 will provide a detailed literature review of the factors affecting the adoption of pedagogies. This chapter will also define key terms that will be used throughout the rest of the thesis. Chapter 4 will present the interview results and detail the causal model developed from these results. This model will show how the pedagogies being used are restricted by standardized tests, the culture of schools, and the content knowledge and pedagogical training of teachers. Chapter 5 will examine the leverage points of this causal model, where enacted policy change could improve the adoption of new pedagogies. Three international case studies will be used in this chapter to illustrate effective of ineffective use of these leverage points. Finally, Chapter 6 will detail policy recommendations that can be used to improve the K-12 STEM Education system.
2. Methods

2.1 Research Design

To develop this model, I have used an in-depth analysis of the literature alongside a comparative case study, focusing on four different case studies in STEM education. Each case focuses on a particular high school. In each case I have interviewed the principal of the school, two STEM teachers who are currently teaching at the school and an administrator from that school’s State Department of Education to determine the decision making factors for each educator. This has led to four interviews from each of the cases, for a total of 16 interviews informing the project. These cases were then compared to one another to identify policies and practices that allow flexibility to adapt to pedagogical barriers. While this study has a small sample size, which has influenced the generalizability of the model, these results were used to direct and apply broader results from the literature. Conducting interviews for each role of educator across different cases allows cross-role comparisons as well. I was able to look for similarities and differences amongst the three different roles represented (administrator, principal and teacher) to bolster the external validity of the results. The different cases were diverse enough to show the range of actions and decision making in the education system.

Using case studies is the most effective methodology for understanding the causal relationships of interest. Case studies are particularly effective when asking causal questions of contemporary issues where one does not have control over the events.\(^\text{15}\) In this instance, observing the real world interactions of educators was of particular interest, so it would be detrimental to use a classical experimental model with treatment and control groups. Such an experiment would have to be widespread across many different schools and states to not simply develop a model specific to that one context. While there are a number of methodologies that deal with events not in the experimenter’s control (i.e. survey or archival analysis), most of these deal with questions of “what?” and “how many?”\(^\text{16}\) While this thesis has been interested in “what works and what doesn’t in STEM education,” the primary focus is on how and why new pedagogies are adopted.

\(^{15}\) Yin, *Case Study Research: Design and Methods, 4th ed.*, (California; Sage Publications, Inc, 2009), 8.

\(^{16}\) Yin, *Case Study Research: Design and Methods, 4th ed.*, 10
This study is inherently limited because of the small sample size. Within these sixteen interviews, there was a wide breadth of opinions expressed, but this sample was by no means exhaustive. There will be numerous schools and specific contexts for which these results, and the resulting recommendations, are not applicable. Case studies are typically criticized for a lack of generalizability to the population as a whole because of the small number of participants. It is true that case studies lack statistical generalizability, where the results can be generalized for the population as a whole. In the case of education, however, this type of generalizability is almost never possible without a sample size of thousands of schools because there is so much variation between schools, districts and states. However, well-selected case studies can have analytic generalizability, where the results are used to advance theories developed from in-depth analysis. The results of this set of case studies have been used to develop an in-depth model of pedagogy use in schools.

The level of analysis used in this study is the individual educator. Though the outcomes of the school as a whole are important, the goal of the study is to understand how each educator makes decisions about pedagogy and interacts with other members of the education system. Therefore, the focus is placed on the individual educator rather than the aggregate school.

Campbell and Stanley have written a well-known text for social science researchers on how to design studies to maximize their validity, both internally and externally. Internal validity is concerned with the ability of researchers to draw well-supported conclusions from their research while external validity is concerned with how well researchers can generalize their findings to a wider context. Especially in education, there is a tradeoff between external and internal validity; to make stronger claims about the successes of specific programs, researchers restrict themselves to only studying one particular school and then are never able to generalize out of their school context. In this situation, where the goal is a causal model for that can explain across multiple schools and guide general decisions, the primary concern is for maximizing the external validity.

---

17 Yin, Case Study Research: Design and Methods, 4th ed., 10
2.2 Case Selection

The American education system is incredibly diverse not only in the quality of education provided and in those who receive it, but also in the school structure and focus of that education. As such, it is not appropriate just to examine one school and draw conclusions about what factors affect STEM education. This work focuses on four different types of public high schools and develops a case from each:

- STEM Magnet School – Magnet schools draw the best students from the district, county or state by reviewing applications based upon previous merit and test scores. They are funded differently than regular public schools and are focused on STEM education.
- STEM-focused School/ Project based school – Distinct from a magnet school, this school educates students who live in the surrounding area and has no admission criteria. Uniquely, this school will have a novel curriculum focused on providing STEM education in innovative ways.
- Comprehensive High School – Similar to a STEM-focused school, a comprehensive high school serves the community it is in and has no admission standards. This model, which represents the vast majority of schools in the United States, does not focus on providing STEM education aside from the standard-offered STEM subjects.
- STEM Turnaround/Transformational School - Under the federal No Child Left Behind Act, persistently low performing schools must undergo a rapid improvement by becoming a) a turnaround school, where the principal and 50% of the staff are replaced; b) a transformational school where rigorous evaluations and standards are put in place; c) closed and reopened as a charter school; or d) closed permanently. Many of the turnaround schools and transformational schools also choose to refocus their efforts to providing high quality STEM education. They have a persistent low-performing student body and often undergo tumultuous change in a short period of time, opening the door for major restructuring.

Each case was selected through a purposive snowball sampling method. I reached out through previous contacts to find administrators responsible for STEM education in their state.

Then I used a variety of metrics to select specific high schools for study within the jurisdiction of these administrators. While initially tempted to select very successful STEM schools, selection of cases based on the dependent variable studied often comes with many problems for analysis and often limits the quality of information that can be used in the analysis.\textsuperscript{20} By limiting the cases to only successful STEM schools, this work would have only identified what factors might have been important to their success. These cases were selected on one independent variable that we are particularly interested in (the amount of STEM-focused learning that takes place) and two independent variables that we expect from our literature review to be most impactful on outcomes (the socio-economic status of the students and the previous academic achievement of these students). We then selected our cases so that they could fulfill the following experimental design:

<table>
<thead>
<tr>
<th>Factor</th>
<th>STEM Magnet</th>
<th>STEM-Focused</th>
<th>STEM Turnaround</th>
<th>Comprehensive HS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socio-Economic Status</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Type of Curriculum</td>
<td>STEM Project Based</td>
<td>STEM Project Based</td>
<td>STEM Project Based</td>
<td>Standard</td>
</tr>
<tr>
<td>Prior Achievement of Students</td>
<td>Admissions Exam Required</td>
<td>No Admissions Exam</td>
<td>No Admissions Exam</td>
<td>No Admissions Exam</td>
</tr>
</tbody>
</table>

In this design, it is possible to compare outcomes and schools across these three factors and try to eliminate the confounding effects from the eventual analysis. Using this method a STEM Magnet school that was particularly successful, a STEM-Focused school that has been marginally successful, a STEM turnaround school that has not yet been successful and a Comprehensive high school that have been particularly successful were selected.

When identifying individual interviewees, administrators were not allowed to refer us to specific schools or teachers within their school. This was especially important when selecting

teachers for the study to remove selection bias, where we only interview teachers that reinforce the same viewpoints as the principles. Instead, I randomly selected STEM teachers to interview at each school.

There are, however, still numerous external factors that could threaten the internal and external validity of our case study. Most of these factors, such as quality of teachers, funding situation, support from surrounding community, etc. are part of the institutional structure that schools operate within and as such we would like to explicitly consider these in our analysis. By not selecting cases to control these variables, it is possible that there are a range of schools that might fall as outliers to the causal relationships developed in the next chapters. There were four cases of experimental mortality, where three teachers and one administrator who were contacted and agreed to be a part of the survey were later unresponsive to repeated email contact.

2.3 Data Collection

To collect this dataset, two instruments were used: a closed-ended survey and an open-ended interview. The close-ended survey was used broadly to categorize the responses of interviewees and give direction for each of the open-ended interviews. Interviewees were asked to respond on a seven-point Likert Scale (Strongly Agree, Moderately Agree, Slightly Agree, Neither Agree nor Disagree, Slightly Disagree, Moderately Disagree, Strongly Disagree) for each statement. The statements were developed from opinions on STEM education issues found throughout our literature review. The survey can be found in Appendix A. These results were then used to guide the interview process, where time was spent examining the reasoning behind particularly strong beliefs on certain issues.

Interviews proceeded in two stages: the administrators and principles in the first group of interviews followed by the teacher interviews. This two-stage design allowed the results of the first interviews with principals and administrators to guide the design of the second interviews with teachers. Analysis of the first results was done to identify causal relationships when administrators and principals attempt to guide teachers in the classroom; the second interview protocol was designed to investigate how teachers react to these specific actions. The interview protocols can be found in Appendices B and C, respectively. The surveys and interviews were first piloted with administrators, principals and teachers that were not to be included in the case
studies. However, these interviews were still informative and will be used to help guide the creation of the causal model.

There has been one difficulty with the data collection process: the response rate for administrators and principals was extremely high, nearly 100%, but the response rate for teachers was much lower. One of the schools selected was not included in the sample because, despite an insightful interview with the principal, none of the teachers responded to interview requests. This principal interview, however, is included in the cross-role comparisons. Overall, this selection effect may impact the internal validity of our work, because there was obviously some self-selection effect on the part of teacher respondents.

2.4 Measurement of Key Variables

The operationalization of variables, where researchers define the fuzzy or unclear variables they are interested in so that they can be measured, was a unique challenge in this study. Many of the causal relationships that we are interested in are not well defined by the educators themselves, let alone for outside observers who are not familiar with that particular school. Thus, not only do researchers have to operationalize our own definitions, but we must also understand those of the educators themselves. The research design in this work is essential in getting to these operational definitions: first the survey results are used to hone in on beliefs that educators hold strongly and then the focused interview questions on these beliefs are used to understand how educators make their decisions.

To better track the decision-making of educators, interviews were coded to make the data easier to analyze. Broadly, interviews were coded as information input into an educator’s decision-making, the factors guiding their decision, and the actions that they can take. Within each of these codes, separate sub-codes were developed as follows:
<table>
<thead>
<tr>
<th><strong>Educator</strong></th>
<th><strong>Inputs into decision-making</strong></th>
<th><strong>Factors guiding their decision</strong></th>
<th><strong>Actions that can be taken</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrators</td>
<td>Standardized test scores</td>
<td>Legislative pressure, overall public opinion</td>
<td>Curriculum setting, testing requirements, teacher salaries</td>
</tr>
<tr>
<td>Principals</td>
<td>Standardized test scores, course grades, college admission statistics,</td>
<td>Parental opinion, curriculum and testing requirements, school culture, ability of teachers</td>
<td>Teacher hiring and firing, prioritization of school resources, professional development</td>
</tr>
<tr>
<td>Teachers</td>
<td>Standardized test scores, regular assessments (both formal and informal),</td>
<td>Parental opinion, curriculum and testing requirements, time available, knowledge of subject, school culture, previous training</td>
<td>Content taught, pedagogy used</td>
</tr>
</tbody>
</table>

Of course, not all responses fit nicely within these codes. New codes were created as they have been appropriate along with an "other" section to capture responses that do not fit well within the codes. Once all the new codes were created, the previously coded responses were reviewed to ensure that the first interviews were coded with the entire set of codes used in the last interview. This ensured that the measures are the same across all interviews.

These codes and operationalized measures are therefore quite reliable. Because all the coding was done by the same coder, the meaning of each code was consistent across each interview that was coded. Since the interviews were revisited with the new codes generated by the coding process, the codes were also consistent across all of the interview data.
2.5 Data Analysis

The analysis of this gathered interview data was aided by the development of our causal model. The model creation has been approached as an exploratory experiment, thus it began with a proposed model based on the literature review, and then it was revised iteratively to support new or contradictory causal loops supported with interview data. Further, gaps in the causal relationships were filled by either revisiting the interview data for any critical data that were missed or by using well-developed relationships in the literature. Only the final model, with its support both from the interviews and literature base, is included.
3. Literature Review

The literature on implementation issues in schools proffers many answers to why new pedagogies have not been widely adopted. In the literature, the success of an attempted implementation depends heavily on that school’s particular context. Two studies of the implementation of similar pedagogies often identify very different reasons for failed reform. This makes it difficult to connect these barriers to change outside of the context they were studied within. However, there are general themes to the barriers encountered:

- There is a lack of incentives for teachers to adopt new pedagogies because of the mismatch between test scores and student knowledge
- The teacher workload impedes the ability of teachers to effectively use the new pedagogy
- The school culture resists changing pedagogical practice
- There is miscommunication between different actors in the education system.

This review will begin by describing how the usage of improved pedagogical methods will improve test scores. Each of these barriers, as identified in the literature, will be explained in detail in this chapter.

3.1 Improved Pedagogies Necessary for Higher Student Achievement

The selection of an appropriate pedagogy is centered on the definition of knowledge that we want to instill in students. If one believes that scientific knowledge is the set of facts, procedures and formulae that scientists use, then one will select an appropriate lecture-based pedagogy which emphasizes the use of rote memorization to memorize these facts. 21 This was the commonly held belief of educators around the turn of the 20th century, and the pedagogies used for the last century have reflected that. However, our understanding of what “knowledge” is has improved drastically. Though there are still some unanswered questions within the cognitive

---

science community, in this thesis we will use the most well supported of these theories. As such, we define deep student knowledge to not just be the series of facts, procedures and formulae that are remembered and applied, but also the strategies of using those facts, how each fact relates to another and the practices that experts use to communicate and explain to others. For example, knowledge of Newton's Laws is not just the wording, the formulae and how to use free body diagrams, but also how these laws can be used in concert with concepts such as friction and torque to build theories about the motion of all macro-scale objects and develop experiments to test these theories.

A modern definition of knowledge also needs modern pedagogies. In keeping with this definition of knowledge, the National Research Council has concluded that "students learn science by actively engaging in the practices of science." Without this active learning, as in the case of rote memorization or other lecture-based pedagogies, students end up with a fragmented and primarily fact-recall understanding of science. Traditional hands-on lab activities are also not enough to spur in-depth knowledge, as they typically focus on data collection and analysis with little treatment of building theories and models or checking them to make sure they are internally consistent. This is not to say that lecture-based pedagogies do not have their place, as they can effectively transmit material in certain situations. However, a student of science cannot develop the necessary skills without also taking part in in-depth investigations.

To develop this type of deep student knowledge, one with deep connections across topics and novel methods of applying facts and formulae, students need to "learn by doing." Research has shown that constructivist pedagogies, where students are able to construct their own understanding of how a phenomenon works, develop knowledge that is "scaffolded" into place by connections to topics that a learner already intuitively understands. Rather than simply being told how something works by a teacher in diagrams on the chalkboard, students are guided by

27 Ibid.
the teacher through a situation where they experience a phenomenon firsthand and experiment with it until they develop their own understanding. Such situations could be active experimentation, game playing, nature walks, and so on. In one account, students learning about pendulum motion “played” and experimented with different arrangements of pendulums for an entire week, getting an intuitive feel before the teacher ever began guiding the students to learn specific topics. This knowledge can be very personal: a football player might root their understanding of parabolic motion in the context of a thrown football and momentum in the context of being tackled by a much larger linesman. This research on the effects of constructivist pedagogies spans the last three decades and has been catalogued meticulously by Duit,\(^{28}\) who concludes that the evidence to support the use of constructivist pedagogies in schools is overwhelming.\(^{29}\) Despite this, teachers are still predominately using the same learning by rote memorization methods that have been used in classrooms for hundreds of years. There are numerous examples of improved pedagogical methods, yet none have received widespread adoption.

One example of an innovative constructivist pedagogy is Project Based Learning (PBL), which assigns students complex problems that students explore at their own pace. Teachers do not guide students step by step through the project; rather, they act as facilitators to help students develop their own understanding. A large number of studies have shown that PBL has been found to both increase long-term retention of concepts and improve student attitudes towards learning.\(^{30}\) One study argues students develop more sophisticated methods of problem solving because they understand the content more deeply.\(^{31}\) Despite this success, PBL is not widely adopted and is used only in at most 0.2% of classrooms nationwide.\(^{32}\)

Another promising constructivist pedagogy is the flipped classroom. Here, students use video-based lectures or interactive text to learn content at home and then are guided by their teacher in the classroom performing exciting activities and problem solving: what is traditionally


\(^{29}\) Treagust and Duit, “Multiple Perspectives on Conceptual Change in Science and the Challenges Ahead,” Journal of Science and Mathematics Education in Southeast Asia, 32(2009).


\(^{31}\) Ibid

\(^{32}\) “Project Lead the Way About Us,” http://www.pltw.org/about-us/who-we-are

27
thought of as homework. However, many educators balk at the time it would take to develop these resources for use at home and rely on online repositories of content. One successful content creator is The Kahn Academy, which develops high quality videos that are meant to be streamed over the internet. *Time* magazine named the founder of Khan Academy, Sal Khan, as one of the 100 most influential people of 2012 for his new pedagogical design. Yet despite the public excitement that Khan Academy has produced, their videos are only used in 15,000 classrooms across the world, at most by 0.04% of the 3.5 million teachers in the United States; not exactly the fundamental change that we all envision. Khan Academy is not unique in having difficulty in adoption and implementation; in fact, it is the norm for such products and ideas in education that attempt to support these new pedagogies.

Students can also improve their learning by teaching that content to another student, known as Learning by Teaching. Most commonly seen in grade schools where an older student will help teach a younger student how to read, this pedagogy has been expanded. Though it takes effort to teach students the proper methods of teaching one another, it has been shown to increase learning gains and make students look forward to learning. In science, this pedagogy can be especially useful for select lessons where students may not have understood the content the first time through, but with more lessons can now derive a deeper content understanding. While the basic uses of Learning by Teaching are commonplace, an advanced usage of this pedagogy is used only in a few select high schools around the country.

With a modern definition of knowledge, we need modern pedagogies to develop that knowledge. This mismatch between the necessity of using new pedagogies in STEM education and the actual usage across the country prompts development of the following four sections, which each offer an answer to the question *Why aren’t high quality pedagogies or technologies being adopted in STEM Education?*

### 3.2 Mismatch between Test Scores and Student Knowledge

One of the most widespread explanations of this lackluster implementation of constructivist pedagogies scrutinizes the effects of standardized tests on pedagogical practice. In

36 Ibid.
a meta-analysis of this work, McNeil found that a large majority of teachers base their decision-making practices on improving student test scores on these exams. One would assume that basing decisions on the feedback teachers receive would be a positive outcome. However, current large scale objective assessments are relying upon the classical definitions of student knowledge and are not accurately assessing the higher order skills of deep student knowledge. The National Academy of Education has found that these standardized tests primarily measure "lower-order" skills; for instance, three quarters of all content on standardized tests was at the "procedural or recall level of cognitive demand." As previously explained, a true test of deep knowledge would include data analysis, reasoning from evidence, making predictions and testing assumptions. In turn, many authors argue, teachers are making poor pedagogical decisions because there is no incentive for teachers to teach deeper knowledge if their students are not tested on them. These teachers can most effectively “teach to the test” by using rote memorization, where students are lectured to and then remember key concepts and formulas.

For example, consider the following two assessments:

Assessment #1
Question: What was the date of the battle of the Spanish Armada?
Answer: 1588 [correct].
Question: What can you tell me about what this meant?
Answer: Not much. It was one of the dates I memorized for the exam. Want to hear the others?

Assessment #2
Question: What was the date of the battle of the Spanish Armada?
Answer: It must have been around 1590.
Question: Why do you say that?
Answer: I know the English began to settle in Virginia just after 1600, not sure of the exact date. They wouldn't have dared start overseas explorations if Spain still had control of the seas. It would take a little while to get expeditions organized, so England must have gained naval supremacy somewhere in the late 1500s.

The second student clearly demonstrated a better understanding of this historical time period than the first student, but most current assessments would mark the first student correct and the second student incorrect. With teachers selecting their pedagogies based on maximizing the results of the Assessment #1, which is primarily a recall-based exam, learning by rote memorization will be instilled in students instead of truly deep student knowledge.

3.3 School Culture Impedes New Pedagogies

Many extensive research projects have also identified the school culture as a driver of teacher behavior and an impediment for reform. Among a number of factors one would expect to negatively affect the implementation of new programs, such as high student/teacher ratio, or a lack of block scheduling, one-on-one tutoring, etc., school culture has been found to be the most significant barrier and has a 2-3 times larger effect size than any other factor. Numerous other studies reinforce the idea of school culture as a driver of school action, and the culture of a school is the primary reason for the unsuccessful adoption of a new pedagogy. Much of the literature on the effectiveness of school principals also focuses on the role that school leaders can play in changing the culture of a school and making it more accepting to reforms. To understand the ability of a school culture to drive actions within the school, we first have to examine the broader organizational culture literature.

There is an extensive body of research describing the effect of an organizational culture on the patterns of behavior and norms of an organization. Van Maanen and Barley first articulated the notion of an “occupational community” as a driving factor behind the workplace

---

actions. In this community, each individual deeply identifies with their occupation. However, when outsiders look at the members of the occupation, they see a homogeneous population among which it is not easy to distinguish one member of an occupation from another. In this way, actions taken by one member of an organization reflect on others. In an attempt to avoid repercussions from the actions of others, each member of the community places social pressure on the others in their occupations to maintain discipline and act a certain way. Van Maanen and Barley go on to describe that the repeated use of social pressure develops a set of normative actions that act as “arguments and accounts to legitimize professional self-control.” A teacher, for instance, does more than simply give a lesson; he educates children. This implies a code of ethics that guides the everyday actions of the educator so as to not reflect poorly on the other teachers at the school. If the teacher will not abide by these norms, he will be ostracized by his peers. The social pressure exerted by an “occupational community” can also be referred to as a social culture, which guides fundamental beliefs, shared values, and behavioral norms of an organization.

This notion of an organization’s culture driving behavior has been well studied in education. As Hoy articulates, for those interested in reforming the actions of school teachers, the intentional guiding or manipulation of the culture of the school will help implement novel pedagogies. Teachers, through many years of practice that have been reinforced by other teachers, have deep and well supported reasons to justify their everyday actions. As new teachers are inducted into the school and taught the rationale behind certain actions, these norms are ingrained into the culture of the school. Any attempt to modify these actions without also changing the normative rationale supporting those actions will be met with stiff resistance. Thus changing the school culture is a necessity in any school reform.

The existing school culture is communicated to incoming teachers via social interactions with practicing members of the community. The internal structure and practices of a school affect the type of culture transmitted to new members. As such, the literature has also identified school-wide practices that are critical to the development of school communities that support the

---

44 Van Maanen and Barley, "Occupational Communities: Culture and Control in Organizations," 317
use of constructivist pedagogies. Louis, Marks and Kruse have investigated exactly this phenomenon, describing how a school culture of professionalism supports teachers in improving their practice. They find that such a professional community has five key aspects: shared norms and values about learning and teaching, a collective focus on student learning, collaboration, deprivatized practice and reflective dialogue. They conclude that, while school culture is not the only aspect that is responsible for school improvement (and that factors such as teacher pedagogical knowledge and content knowledge have an impact), the culture of a school supports these teachers in their efforts to do their job.

Thus while teachers may be alone inside their classrooms, their actions in the classroom could have large repercussions on the ability of other teachers in the school to do their job. Poor performance could lead to increased oversight over teachers, increased pressure from parents, or a restructuring of the school. As a result, teachers put social pressure on other teachers to abide by certain norms of practice, which creates a culture inside the school that dictates the everyday pedagogies that are used.

3.4 Teacher Workload Impedes Use of New Pedagogies

Numerous attempts at adopting new pedagogies have also failed because of the high demands on teacher time. New constructivist pedagogies typically take more preparation time for each hour of class time, which can impede the proper implementation of the reform. As such, upon review of many attempted pedagogies in the last two decades, Cooper has found that there are a “long list of innovations that work only if large increases in resources or teacher time are required.” One study found that the teacher workload can have, in some circumstances, nearly as large an effect as culture on successful implementation. The teacher workload is especially difficult to handle because small fluctuations in teaching time have resounding impacts on prep time. For instance, taking conservative estimates from the literature, if the typical teacher works 12 hours a day during the school year, typically half of that will be spent in

front of the class teaching (6 hours per day). Each hour of teaching requires a minimum of 30 minutes of administrative and grading work each day, and thus a teacher will only have 3 hours, or 30 minutes per hour of teaching, to prepare for class the next day.

This is of particular note to the United States because of the incredibly high workload that the average teacher is given. In the United States, high school teachers are responsible on average for teaching 1051 hours in front of the classroom per year. As seen in Figure 1, this workload is much higher than the average across all OECD countries. To put this in comparison, working the same number of overall hours, Japanese high school teachers have only 500 hours of teaching time in front of a classroom, with a ratio of nearly 3 hours of preparation and administrative work to each hour of teaching. Therefore, many studies have identified that time-intensive pedagogies are not well adopted in the United States because teachers simply don’t have the time to implement new pedagogies.

Figure 1: Time spent Teaching Across OECD Countries

---

52 Data from OECD, *Education at a Glance, Indicator D4*, (2011), can be found in Appendix E.
3.5 Systems Level Miscommunications

Finally, a small body of literature has also identified incoherence and miscommunication across levels of the system as a major barrier to change. While many structural elements of the education system are resistant to changes, unified efforts from teachers, principals and administrators have been successful in implementing new pedagogies. The difficulty lies in aligning the interests of each of these educators across the education system. Top-down approaches fizzle when opposed with stiff resistance by teachers. Bottom-up approaches struggle to gain traction or legitimacy. The decentralized approach to American education, where each state and district within those states has a different approach to education, also leads to situations where educators pursue diametrically opposing actions on education simultaneously. For this reason, the National Academy of Education has concluded that there must be a coherent approach to education reform across all levels of the education system.

In one of the most comprehensive meta-analyses of structural change in education, Elmore concludes that researchers must probe underneath the structures to discover, both conceptually and empirically, what are driving actions in the classroom and how these actions are influenced by principals, administrators and structural factors. Changing one piece without a systematic understanding will often create unintended negative consequences. For this reason, we will develop an understanding in the next chapter of how each of these factors relate to one another, developing a model to inform what policies can best be used to influence the adoption of new pedagogies.

3.6 Glossary of Key Terms

From this literature review, as well as other literature sources that will be used in the following chapters, I have developed a glossary of critical terms that will be used throughout the remainder of the thesis.

Rote-based Student Knowledge - The set of facts, procedures and formulae that students memorize and can use to solve problems. This understanding is often fragmented and is primarily a fact-recall understanding of science.

---

Deep Student Knowledge – More than just the series of facts, procedures and formulae that are remembered and applied by students, deep student knowledge includes the strategies of using those facts, how each fact relates to another and the practices that experts use to communicate and explain to others.

Lecture-based Pedagogy – Teaching practices where the teacher transmits information that students are expected to memorize and understand; typically leads to a rote-based student knowledge (see above).

Constructivist Pedagogy – Teacher Practices where students are guided by the teacher through a situation where they experience a phenomenon firsthand and experiment with it until they construct their own understanding of how a phenomenon works; typically leads to a deep student knowledge (see above).

Pre-service Teacher Education – Teacher training that happens before teachers move into a classroom, typically conducted through a university degree.

In-service Professional Development – Teacher training given throughout a teacher’s career in schools in the form of mentorship, regular professional development days, and other learning opportunities

School Culture – The everyday actions of a school organization that over time become pervasive and self-reinforcing norms, directly influencing teacher behavior

Subjective Assessment – Traditional in-class exams, quizzes, projects or other activities, which are prepared and graded by the teacher themselves. The vast majority of all assessments that students are given are subjective assessments.

Objective Assessment – Assessments that are designed and graded external to the school to provide unbiased data to administrators (i.e. Standardized Tests).
4. Results and Model Development

This chapter will report the interview and survey results, and from these results develop a causal model that explains the continued usage of lecture-based pedagogies in schools. Initially, the results of the surveys and interviews of teachers and administrators offered different perspectives. Teachers reported that, under high pressures to improve test scores, they chose their pedagogy based on a goal of maximizing scores rather than improving the ideal student knowledge as previously defined. Administrators and principals, however, reported that lecture-based pedagogical practice among teachers was often a result of inadequate training and a lack of content knowledge. They argued that standardized tests do not directly influence the adoption of particular pedagogies, but simply reveal poor student performance and put pressure on teachers to improve their pedagogies. To reconcile this discrepancy and investigate the impact of other factors identified in the literature review, I used interview results to construct a causal model of the factors impacting pedagogy in the classroom. In short, the findings are:

- Despite administrators believing that pressures from test scores will drive teachers to adopt constructivist pedagogies, teachers use lecture-based pedagogies to directly improve test scores
- School culture reinforces the pedagogies used by teachers, regardless of their effectiveness
- Lack of in-depth content knowledge leads some teachers to avoid uncontrolled classroom settings
- Constantly overwhelmed for time, teachers rely on pedagogies familiar to them, and teachers without deep STEM training use the methods they are most familiar with, namely lecture-based pedagogies.

I will begin by presenting the contradictory results that motivated the development of the model and then continue by presenting the results of the interviews and survey in each section. Each section will also include the complimentary literature sources that support each model. The full survey results can be found in Appendix D.

4.1 Contradictory Survey Results

Interviewed administrators, principals and teachers held very different views on why constructivist pedagogies are not being adopted widely. Administrators said that teachers needed
more training and content knowledge to utilize new pedagogical techniques, arguing that standardized test scores were not pushing teachers to use lecture-based pedagogical methods. One administrator reported that “teachers are supposed to be teaching the content on the exam regardless of whether or not we are testing them on it. These tests are a way of measuring progress and identifying schools that need additional support.”\(^{56}\) Another administrator acknowledged that “yes, the tests could be improved. But they are designed rigorously and offer insight into how each school is performing; data that are invaluable for the states and districts.”\(^{57}\) One administrator was adamant that these standardized test results focused on providing additional training, saying that “these tests have high stakes, but we only use them to see where and how to place our resources... We also use them to evaluate how and why new training is working.”\(^{58}\) Three of the four administrators surveyed disagreed with the statement that \textit{standardized tests evaluate algorithmic steps or facts, not critical thinking skills}; the fourth administrator only slightly agreed.

Interviewed teachers, on the other hand, almost universally opposed the current application of standardized testing. All eight teachers surveyed agreed with the statement \textit{standardized tests evaluate algorithmic steps or facts, not critical thinking skills}. Most teachers also argued that, rather than providing useful information, standardized tests applied pressure on them to improve test scores.\(^{59}\) One teacher reported that the “consideration of how to increase test scores most effectively is a major factor when choosing how to teach a particular topic. My job essentially depends on it.”\(^{60}\)

So, while administrators believe that standardized tests galvanize more training to improve pedagogy, teachers feel forced to use lecture-based pedagogies that will improve test scores. These are, of course, not the only factors that guide pedagogical decision-making. To explain these seemingly contradictory results, I developed a causal model of the factors impeding progress towards new pedagogies. I will begin by examining the rationale of administrators driving assessment-based accountability schemes.

\(^{56}\) Administrator 4 \\
\(^{57}\) Administrator 3 \\
\(^{58}\) Administrator 1 \\
\(^{59}\) All Teachers except Teacher 5 and Teacher 6 \\
\(^{60}\) Teacher 7
4.2 Administrator Rationale behind Assessment-Based Accountability

Administrators and principals, unable to directly affect change in pedagogy, are forced to rely on pressures from students, parents and standardized tests. Direct actions by principals, such as negative year-end reviews for using lecture-based pedagogies, are reported by principals interviewed as “only putting pressure once or twice a year for teachers to change their pedagogy as opposed to the everyday demands of the job year-round.” This confirms the findings of Marzano, Waters and McNulty, who argue that principals and administrators can only dictate what topics teachers are teaching, not how a teacher teaches that topic. Administrators and principals further reported that it would be inappropriate for them to be involved in the day-to-day actions of a teacher. One administrator said “as a former teacher, I would be uncomfortable telling a teacher how to teach. They know their students best... It is my job though to make sure students are receiving a quality education.”

Without the ability to guide pedagogies themselves, administrators and principals rely on objective assessments to identify where to target pressure for improvement. Traditional in-class exams or final projects are known as subjective assessments and cannot be easily trusted by administrators to evaluate educational outcomes because teachers prepare the exam and grade the results themselves. On the other hand, objective standardized tests provide unbiased data to administrators because of the impartial grading mechanism. “We need to format these tests the way we do to grade them effectively. Having high stakes attached to a test that each teacher grades themselves, and differently from other teachers, would never be effective,” one administrator commented. A principal agreed, saying that “these tests could certainly be improved, but they do provide a rough idea of how each school is doing. It is impossible to use any other type of data to hold schools accountable without biasing the results.”

The results of the standardized tests put pressure on teachers to improve test scores. By attaching high stakes to “failing” such standardized tests, as is done under the No Child Left Behind Act, teachers must pay particular attention to the lowest performing students. Teachers of

---

61 Principal 4
63 Administrator 2
64 National Research Council, Knowing what students know: The science and design of educational assessment.
65 Administrator 3
66 Principal 3
low-performing students do not typically have the same pressures from parents or students to improve. These accountability systems therefore are an attempt to spur action amongst the teachers of these students. Teachers in the Turnaround School certainly reported that this was the case, saying that "test scores absolutely matter to these students. And if they cannot pass the tests, which are not impossibly high bars, we should not be sending them out into the world without the ability to read or write." 

High-performing students also value continued high test scores on national tests, such as the SATs or AP exams, which determine their future college admissions. One teacher felt that "students did not actually care if they learned the content or not, they just wanted to do enough to get an A+ in the class." One teacher described that internally graded measures carried little weight and they could "show a student's performance through high quality projects or internal exams, but all [students] care about is the AP test at the end of the year." Placing an official score on a particular performance or assessment makes it meaningful to students and parents, who then apply pressure to teachers to ensure good performance on these exams. Hanushek and Raymond have also observed the same phenomenon, where negative school report cards have no impact on student achievement while externally-graded standardized tests drive pressure by students and parents.

Administrators believe that these pressures on teachers from low standardized test scores should motivate them to adopt deep pedagogies. One administrator explained the rationale by saying "the vast majority of teachers work incredibly hard. They are doing all that they can every day. Low test scores shouldn't make them lecture harder; the only thing teachers can do is change their practice." Another administrator agreed, saying that "under enough pressure, teachers will change the pedagogies they use. These types of [constructivist] pedagogies take more effort to use, and without any accountability measures pushing teachers to improve the methods they use to teach we won't see any changes." In case studies conducted by researchers,

---

68 Teacher 8
69 Teacher 5
70 Teacher 2
72 Administrator 3
73 Administrator 4
principals or administrators routinely try to use test scores to motivate changes in pedagogies. Another administrator clarified this point, saying that "I know that low student achievement on these tests might not motivate all teachers to change, but it might motivate enough to make a difference for a lot of students." Thus, according to the rationale of administrators, teachers will be able to alleviate these pressures to perform well on exams by using more effective pedagogies.

The rationale of administrators for using standardized tests to drive the pedagogy in classrooms can be summarized in Figures 2 and 3. Once administrators introduce standardized tests, teachers will be able to observe a gap on assessment performance from the desired student performance. In turn, students, parents and schools will strive to raise test scores and fill this gap by increasing the pressure on teachers to improve performance, which is assumed by administrators to push these teachers to adopt new pedagogies or improve the quality of teaching methods. In turn, increased quality of teaching methods, as articulated in the literature review will increase a student's knowledge. With greater student knowledge, the student's performance revealed through assessment, or test scores, will increase, thus closing the gap in assessment performance previously observed. As administrators rationalize it, this interaction should be a balancing loop, where pressure for improving pedagogy arises when student performance is not at the desired level, as seen in Figure 2.

---

75 Administrator 1
Figure 2: Rationale for Assessment Based Accountability

Figure 3 describes how the long-term reliance on test results by administrators and subsequent push towards objective tests has only increased pressure on teachers to improve test scores. As a gap between desired knowledge and current knowledge for graduates develops, society believes that schools are not fulfilling their role and must be held accountable. To hold these schools accountable, administrators rely on the “alleged” objectivity of measures for each teacher, which subsequently will increase the emphasis placed on test results by administrators because the tests are believed to be fair and unbiased. With more emphasis on test results, there is again greater pressure on teachers to improve performance which administrators believe should increase the quality of teaching methods. If these teachers adopt improved pedagogical methods, as administrators assume, students will be exposed to deeper learning and increase a student's knowledge. After a long delay, these students will age and move into the workforce. Some of these students will graduate with a STEM degree and improve the average graduate's knowledge, eventually closing the gap in STEM workforce.
4.3 Shortcut to Improve Test Scores

The previous section detailed how Administrators have rationalized the current system of test-based accountability by assuming that teachers under high enough pressure will adopt deeper pedagogical methods. This section will demonstrate how teachers responded to these pressures to increase test scores by instead using lecture-based pedagogies. However, as seen in the literature review, standardized test scores are not necessarily representative of true student knowledge. Current standardized tests are built upon outdated models of cognition; “classical” theories that view knowledge as a collection of facts, figures, procedures and processes that a learner can use to solve a problem. Teachers interviewed were also cognizant of this, one reporting that “increases in test scores measure exactly that; increasing test scores. They tell you little about what a student actually knows.”

With these tests evaluating only the basic facts and procedures of scientific knowledge, teachers have started to adopt methods to “game” the assessment, rather than focusing on improving quality of teaching. The most common pedagogical approach to maximize these

---

76 Differential and Behaviorist Theories, as articulated in National Research Council, *Knowing what students know: The science and design of educational assessment.*
77 Teacher 7
scores is using rote memorization, where teachers lecture and then students spend time memorizing definitions and repeatedly solving the exact problem types that they will be tested on. One teacher reported that for them “[using rote memorization] is simply is the best way to raise test scores. And it’s discouraging, but that’s what I value most in a pedagogy.” 78 Another teacher reported that this narrowing of pedagogies to lecture-based pedagogies was a direct result of student and parent pressures; “some students and parents actively oppose new pedagogies in favor of pedagogical methods that will increase test scores” 79 These teachers also spent a large amount of instructional time teaching test-specific techniques in the hope of giving students enough “tips and tricks” to do well on the exam, even if students don’t necessarily understand the content. This trend corroborates the literature, as Herman and Golan have found that teachers spend an average of 3-4 weeks of instruction time on standardized test preparation methods. 80

This shortcut to improving test scores without improving deeper student knowledge can be seen in Figure 4. As the pressure on teachers to improve performance increases, teachers increase the time spent on methods to game test scores (rote memorization) rather than increasing the quality of their teaching methods. This increased use of rote memorization increases a student’s performance revealed through assessment, which in turn lowers the gap on assessment performance alleviating pressure on teachers temporarily.

---

78 Teacher 8
79 Teacher 2
Figure 4: Shortcut to improve test scores

While lecture based pedagogies, such as rote-memorization, are able to increase standardized test scores, teaching by rote is proven to be detrimental to deeper student knowledge. Hiebert and Carpenter,81 Tobin,82 and many other education researchers have found that procedures and facts learned by rote are easily forgotten and prone to errors. The largest gains in test scores via rote memorization are through “cramming” for exams, relying on short term memory to be able to “parrot back answers” the next day on the test. The continued usage of rote memorization will lead, as discussed in the literature review, to a rote-based student knowledge composed of fact-recall instead of on a deep understanding of the content.83 Thus, when exposed to incredibly high pressures to improve test scores which are poorly designed to measure deep student knowledge, teachers end up relying upon the least valuable methods of learning (lecture-based pedagogies) in an attempt to improve not the deeper student knowledge but the student test scores.

---

83 National Research Council, Knowing what students know: The science and design of educational assessment.
The use of this shortcut will troublingly lead to long-term negative unintended consequences across the entire system, as seen in Figure 5. As more teachers teach to the test using rote memorization, test scores might rise but deep student knowledge will not increase. Without improvements in deep student knowledge, average graduate knowledge will not improve and the gap between desired knowledge and expected knowledge will persist. This will put even more pressure on teachers to improve performance. This reinforcing cycle will lead to a continued decline of student knowledge as the pressure on teachers continues to encourage them to use rote memorization to improve test scores.

4.4 Culture Reinforcing Teaching to the Test

School culture, previously defined as the everyday actions of a school organization that over time become pervasive and self-reinforcing norms, is incredibly influential to the everyday actions of teachers. These cultural norms transmitted between teachers determine acceptable or unacceptable behavior in the school which, among numerous other impacts, drives the adoption of pedagogies. Some school cultures narrow teaching down to lecture based pedagogies; others discourage rote memorization and other such lecture based pedagogies in

---

Figure 5: Long Term Effects of Rote Memorization

---

favor of constructivist pedagogies.\textsuperscript{85} The schools examined in this study confirmed the ability of culture in driving pedagogical methods.

Of the four schools examined, three had cultures which emphasized acquiescing to pressures from parents to use lecture-based pedagogies. One teacher reported that “at first I was hesitant to share what I was doing in my classroom [partly due to being ashamed], but other teachers were using the exact same methods I was. It helped to have a peer group.”\textsuperscript{86} Five out of the eight teachers reported regularly using lecture-based pedagogies in their classroom, many reporting that “while there may be other methods, at our school we think that lecturing works well for our students.” It was also observed that teachers who do not follow these norms are quickly isolated by other teachers, as was claimed in the literature.\textsuperscript{87} This culture was especially pervasive at the Turnaround School, where one teacher remarked “our core mission is to serve students, and we cannot do that if our school is closed. We, as a school, stick together to raise test scores.”\textsuperscript{88}

Culture was also behind one school’s resistance to lecturing. At the STEM Magnet School, one teacher remarked that “I can experiment with new ideas because everyone knows I am a professional. The [school] administration treats us as professionals as well.”\textsuperscript{89} The same teacher also remarked that their school culture had changed the pressures on students so that they “understand that a great [recommendation] letter can do more for them than changing that A to an A+. With this in mind students are more willing to buy into new pedagogical approaches and participate above and beyond what is expected to distinguish themselves to me.”\textsuperscript{90} This principal was able to effectively guide his teachers’ actions by “giving them authority over their classroom, giving them resources, then expecting excellence,” which confirms the approach shown to be successful in some literature sources.\textsuperscript{91}

The effect of a school culture on teacher practice is magnified by the current method of pre-service and in-service teacher training. Much of this training process happens in the school

\textsuperscript{85} Sarason, \textit{The Predictable Failure of Educational Reform: Can We Change Course before It's Too Late?} (Jossey-Bass, Inc.: San Francisco, 1990).
\textsuperscript{86} Teacher 8
\textsuperscript{87} Louis, Marks and Kruse, “Teachers' Professional Community in Restructuring Schools,” (1996).
\textsuperscript{88} Teacher 7
\textsuperscript{89} Teacher 4
\textsuperscript{90} Teacher 4
context, both through internships (ranging from a month to a year-long) and in the first few years of teaching. The pedagogies learned during this time are embedded in teachers for the rest of their careers. Thus, when a school culture emphasizes the use of lecture-based pedagogies, budding teachers who take their practical training in this environment will tend to use these pedagogies for the rest of their careers. One teacher remarked that “the new teachers come in the door with little idea of how to teach effectively. We have a mentorship program that pairs an experienced teacher with them to show them the ropes. Sometimes what I suggest is much different than what they learned in their classes, but they soon see the value.” One new teacher remarked that “my first months of teaching were incredibly hectic, so when my mentor provided any resources at all, those were the ones I used.” Often, the method passed along is rote memorization. In three of the four case study schools, the current culture of the school encouraged rote memorization through both norms that actively encourages the use of the practice and passively has no practical knowledge of other pedagogical methods.

The reinforcing effect of a school’s culture can be illustrated in Figure 6. As the time spent on methods to game test scores (rote memorization) increases, the acceptance in school culture of teaching to the test will also increase, thus reinforcing the use of the pedagogy. Once the school culture emphasizes teaching to the test, the culture will be forced on other teachers and the overall ratio of teachers teaching to the test in a school will also increase. With a greater number of teachers using these methods, the time spent methods to game test scores (rote memorization) will increase, leading to a reinforcing cycle. It is important to note that, while currently many schools are locked in a reinforcing cycle, the same loops could encourage the use of new pedagogies. If the only acceptable pedagogies are constructivist, other teachers will face social pressure to use similar methods, as is the case in the STEM Magnet school.

94 Teacher 6
95 Teacher 2
4.5 The Effect of Content Knowledge on Pedagogy

It is a commonly held belief among educators that teachers with deep content expertise produce better learning results. The literature, however, reveals no direct correlation between advanced coursework among teachers and student outcomes on assessments. While many studies have found significant gains in achievement from a teacher’s advanced content knowledge, other studies have found only small gains or none at all. Other research explains these findings by suggesting that there is actually a threshold effect for content expertise: the effects of more than a year or two of a teacher’s intense study in content specific courses becomes negligible for student outcomes. High-school courses draw upon, at most, the

---

first few classes of a teacher’s undergraduate study, and more depth of content knowledge does not contribute significantly to their ability to teach high-school math or science. However, these interview results suggest two paths for how content knowledge can influence pedagogical use in ways not well treated in the literature.

Administrators and principals reported that many teachers were uncomfortable in uncontrolled classroom settings because they did not have an in-depth knowledge of the content they were teaching. One principal described this phenomenon by saying that “so many teachers are scared to have the wrong answer, especially in an honors classroom where the students are particularly insightful. Such classes can quickly outpace the knowledge of teachers.” As such, teachers without deep content knowledge tend to stay away from pedagogies such as project-based learning, where teachers allow students to explore content on their own and often stray into new and uncharted content. For instance, one teacher articulated his initial hesitation of using such constructivist pedagogies saying “it can be a bit awkward at first, saying ‘I don’t know’ in a classroom. Students expect us [teachers] to have all of the answers, and it can undermine your ability to be the expert in the room. But it is actually all about classroom culture, something that takes a lot of work to develop, where you can say those big three words and keep legitimacy.” Other teachers agreed in the interviews, typically trying to avoid such situations where they would, as one teacher articulated, “lose control of the classroom.”

Teachers without deep content knowledge also spend preparation time learning the content and relying on pre-made materials rather than researching and developing effective pedagogical methods. A few teachers reluctantly reported that there were many units or even courses where their content knowledge was at a very basic level. To teach this content effectively, they had to first spend time learning the content themselves before planning the lessons. One teacher commented that “when the content is difficult for me, all of my time gets spent making sure I can answer direct questions about the material. In turn, I often end up relying on pre-made materials. If these materials facilitated project-based learning, I would do that…but they are

---

102 Wilson, Floden, & Ferrini-Mundy, *Teacher Preparation Research: Current Knowledge, Gaps and Recommendations.*
103 Principal 1
104 Teacher 8
usually lecture-based materials." This will be an important leverage point for improving constructivist pedagogical adoption, as developed in the next chapter.

This survey and interview data suggests that deep teacher content knowledge allows teachers to use more effective pedagogies. This effect can be seen in Figure 7, where an aspiring teacher will gain a lecture-based knowledge as a student, which will be retained once they graduate. This lecture-based knowledge will then guide the pedagogies once these graduates become teachers. In turn, this teacher workforce will not be successful at passing knowledge on to others because of the little content expertise they can draw upon or an unwillingness to experiment with new pedagogies, leading to a reinforcing loop. This also implies that the opposite direction of the loop should also be true, where those teachers who are taught by constructivist pedagogies will be able to pass along a deep understanding to their students.

Figure 7: The Effect of Content Knowledge

4.6 The Effect of Pedagogical Knowledge

Teachers rely on the same pedagogical methods that they learned from, which cannot be changed without time for pedagogical training and experimentation. As discussed above, the previous content knowledge of a teacher dictates the level of pedagogical depth they can explore.
with their students. Marks has found the same to be true of pedagogical technique: teachers have difficulty conceiving of new ways to teach a topic outside of the method they use to understand it.  

This type of emulation between teachers and their former teacher is typically referred to as mimetic isomorphism, creating a legacy effect where pedagogies are propagated forward not because of their merit, but because they reflect how teachers understand a topic. In this way, Eisenhart et al. have shown that, without strong pedagogical training, teachers focus on simply describing their own understanding of a topic to students, assuming that the students can understand the content the same way. Little time is therefore spent developing student understanding of the topic by experimenting with various pedagogical approaches until they find the "right one." 

Time pressures on teachers exacerbate this reliance on the same "legacy" pedagogies. As discussed in the literature review, teachers in the United States carry one of the heaviest teaching workloads in world, which limits the amount of time they can spend on pedagogical preparation. One principal described this situation in which, "teachers are overwhelmed all the time. Without firm pedagogical backing teachers will just find any life raft they can to keep their head above water. Unfortunately this is often teaching to the test." All teachers interviewed commented that they would like to teach using constructivist pedagogies, but the teaching demands placed upon them limited the amount of time they were able to put into preparation. One teacher said that "I often have great ideas for new methods I would like to try, but they simply get lost in the grading and prepping for the next day."

The effect of pedagogical knowledge on pedagogical practice is therefore similar to the effects of content knowledge, as shown in Figure 7. The pedagogical knowledge of teachers influences the quality of teaching methods used, which will lead to increases a student's knowledge. After a long delay, these students will become graduates, some of whom will move

---

109 Ibid.
111 Principal 3
112 Teacher 1
into the teaching workforce, where they will rely upon their knowledge and previous training to teach students.

**Pedagogical Knowledge for Teachers**

![Diagram of Pedagogical Knowledge for Teachers]

**Figure 8: Effect of Previous Training on Teaching Ability**

4.7 **Full Causal Model of Factors that Effect Pedagogy**

It is clear after examining the literature and interview data that there is no one factor that determines whether a new pedagogy will be adopted; rather, there is a complex system of factors that all influence each other. To summarize:

- Pressures from test scores drive teachers to use lecture-based pedagogies.
  - Test scores can be improved by methods other than just increasing deep student knowledge. Instead of alleviating pressures by improving their pedagogy, teachers can improve test scores directly through lecture-based pedagogies, inadvertently hurting students in the long run as they fail to develop real student knowledge.
• School culture reinforces the pedagogies used by their teachers.
  o This culture can either push teachers towards using lecture-based pedagogies or
towards constructivist pedagogies.
• Lack of deep content knowledge limits the pedagogies used.
  o Teachers sometimes feel uncomfortable in uncontrolled classroom settings because of
  their lack of content knowledge.
  o Teachers without deep content knowledge spend preparation time learning the content
  and relying on pre-made materials rather than using new or more effective
  pedagogical methods.
• Lack of pedagogical training limits the pedagogies used.
  o When overwhelmed for time, teachers rely on pedagogies familiar to them.

The next chapter will identify the key leverage points within this system where policy changes
could be made to improve the adoption of new pedagogies. I will then identify how effectively
other countries have used these practices to improve their teachers’ pedagogical methods.
5. Leverage Points and International Comparisons

In the previous chapter, a plausible causal model illustrated how the use of poor pedagogies is incentivized and reinforced in schools. This chapter identifies the leverage points in the system; points where changes in policy could encourage the use of more effective pedagogies. Before developing more general policy recommendations in the next chapter, I will offer potential solutions that take advantage of each of these leverage points. I will then examine three cases of educational reform in other countries to observe how some of these solutions were either used effectively or ineffectively. In short, these leverage points and their potential solutions are:

- **Leverage Point: Assessment Design and Objectivity of Measures**
  - Solution 1: Remove Standardized Tests
  - Solution 2: Improve the Quality of Standardized Tests
  - Solution 3: Legitimize the Results of Subjective Assessments

- **Leverage Point: Teacher Pedagogical Knowledge**
  - Solution 4: Stronger System of Pre-service Education
  - Solution 5: Develop a Robust System of In-service Teacher Professional Development

- **Leverage Point: Teacher Content Knowledge**
  - Solution 6: More Effective STEM Training in Undergraduate Education
  - Solution 7: Bring Expert Materials and Knowledge into the Classroom

- **Leverage Point: Culture and Competition**
  - Solution 8: Use Pre-service Clinical Training Programs to Ingrain a Particular Culture into Teachers

5.1 Leverage Points

5.1.1 Assessment Design and Objectivity of Measures

One of the critical problems identified through the interviews and survey results is the use of lecture-based pedagogies, like rote memorization, by teachers under pressure to improve test scores. As developed in the literature review and the previous chapter, there is a mismatch between true student knowledge and the current assessments. Current standardized tests
primarily measure “lower-order” skills; researchers have found that three quarters of all content on standardized tests was at the “procedural or recall level of cognitive demand.” An assessment of true student knowledge would probe the connections and applications of content knowledge by including data analysis, reasoning from evidence, making predictions and testing assumptions. With such immense pressure on teachers to improve the standardized tests scores that do not represent deep student knowledge, there is little incentive to use constructivist pedagogies instead of simply trying to maximize test scores via lecture-based pedagogies.

There are three fundamental methods of solving this problem: removing standardized tests, adopting a method of grading subjective tests that legitimizes their results, or developing objective assessments that are harder to “game” by using rote memorization to increase test scores without improving student knowledge.

Solution 1: Remove Standardized Tests

Removing standardized tests is the simplest solution. As shown by McNeil, this solution would allow teachers the freedom to experiment with constructivist pedagogies.\(^{114}\) However, there is still a perceived need to hold schools accountable and ensure a high quality of education for all public school children. As such, administrators would still require an accountability system to identify ineffective teachers and either put pressure on them to improve their practices or terminate their contract. This system of accountability for teachers and schools should be based upon qualitative assessments of performance, such as principal assessments and regular in-class observations from experts. Rather than using improved student test scores as a rough measure of pedagogical practice, principals and administrators should be directly evaluating teachers on their pedagogical approaches during lesson observations.

Assessments of student knowledge could be a blend of both objective and subjective measures to provide a more accurate picture of student achievement. For instance, basing a student’s classroom grade on a portfolio filled with many end-of-semester projects, well-written essays, and teacher evaluations of deeper skills such as critical thinking and persistence could be used to provide a detailed account of deep student knowledge to students, parents, college

---

114 As seen in McNeil, *Contradiction of School Reform: Educational Costs of Standardized Testing.*
admissions offices and so on. This portfolio approach is already used in up to 20% of college admissions, where objective test scores from the SAT or AP exams are either not considered or hold relatively little weight in the admission process. Adopting a portfolio approach would thus remove pressures from parents or students to directly raise test scores, and would instead incentivize students to raise the quality of their portfolio by producing higher quality work.

**Solution 2: Improve the Quality of Standardized Tests**

The second method to circumvent the use of lecture-based pedagogies to improve test scores is to enhance the quality of the assessments themselves. If standardized tests assess more than simply lists of facts or applications of specific theories, the only reliable pedagogies to improve test scores are those constructivist pedagogies that increase true student knowledge.

For example, a test might be able to ask:

- **Question:** Which material is more dense; tin or gold (give their weight and volume)?
- **Question:** Would these metals float in water? How could we know this by knowing their density?
- **Question:** If you use tinfoil (very thin tin) to make a boat, why does it float?
- **Question:** Now you crumple up the tinfoil and it still floats in water. How would this change your previous answer (if at all)?

New assessments can use follow-up questions to better evaluate a student’s deeper knowledge and how they transfer knowledge to a novel situation. A student who learned this content via rote memorization, for instance, would only be able to answer the first question and perhaps the second question on this improved assessment, giving a student at most $1/2$ of the correct answer. Only students who have experimented and deeply understand the concept of density will be able to get all four questions correct, incentivizing teachers to select pedagogies to develop that deep understanding of the concept.

The push from administrators for more objective testing methods to eliminate any sources of bias or cheating, however, limits the quality of assessments that can be given. Current

---


methods of observations in large-scale assessments use a multiple-choice scantron that can be fed into an automated grading machine to quickly and cheaply grade exams. Though some of these multiple choice, true or false, and fill in the blank questions can be quite sophisticated, reliance on these techniques narrows the range of student knowledge that can be assessed to primarily recall-based or process-based answers.\textsuperscript{118, 119} As noted by the National Research Council, critical thinking and problem solving abilities cannot be easily assessed by filling in a blank or choosing true or false.\textsuperscript{120}

New technologies and methods could provide more accurate information about each student. Computer models can be used, for instance, to allow students to manipulate a model of an electrical circuit to troubleshoot and then suggest ways to fix the problem.\textsuperscript{121} Alternatively, researchers can use low cost cameras to track students’ eye movements while taking assessments. These cameras measure length of gaze and the sequence of information considered to assess how the student is interpreting information and the strategy they are using to solve the problem.\textsuperscript{122} Performance tasks, where students are given a set of supplies and asked to complete a problem by manipulating physical objects, could also be used to assess a student’s real world understanding of a phenomenon outside of simply memorizing the process and result from a textbook. These new techniques, however, are still in their infancy, and may not be ready to be used in high-quality assessments.\textsuperscript{123}

Solution 3: Legitimize the Results of Subjective Assessments

Subjective assessments that are graded by the teachers themselves can ask questions that measure deep student knowledge more accurately than current standardized tests. Again, basic multiple choice, fill in the blank or true and false questions can only measure a narrow range of a student’s understanding of a topic, such as fact recall.\textsuperscript{124} Subjective exams, however, are best graded by the teacher themselves, which leaves many administrators to believe that these

\textsuperscript{118} National Research Council, \textit{Knowing what students know: The science and design of educational assessment.}
\textsuperscript{120} National Research Council, \textit{Knowing what students know: The science and design of educational assessment.}
\textsuperscript{121} Partnership for Assessment of Readiness for College and Careers.
\textsuperscript{123} Whiteboard Advisors, \textit{Tracking Measures, Common Core Materials, and Other Timely Topics in Education}, (2013).
\textsuperscript{124} National Research Council, \textit{Knowing what students know: The science and design of educational assessment.}
assessments are open to bias. By implementing a subjective grading system that eliminates most bias from the system, students could be assessed with questions that probe the deeper understanding critical to true student knowledge.

There are proven methods for removing subjectivity with teacher-peer grading that could reliably be implemented across the country. Perhaps the most robust of these grading systems was developed by the International Baccalaureate (IB) program and has since been adopted worldwide. To score these external exams, graders are chosen from among IB teachers in high schools and grade assessments from around the world, ensuring an unbiased review of student work. These teachers then send a small sample of exams to more senior graders who ensure accuracy and consistency among the graders from all over the world. These exams are administered to over 1.1 million students in 145 different countries, proving that such systems can maintain objectivity at a large scale.

5.1.2 Teacher Pedagogical Knowledge

Even if educators could break the use of lecture-based pedagogies by aligning assessments more closely with the desired student knowledge or getting rid of assignments all together, this complex system would still be hampered by the various reinforcing loops in the causal diagram. Teachers, without firm pedagogical knowledge and lacking time to learn new pedagogical approaches, would often still resort to the same rote-based pedagogies they used to learn that content. To solve these problems, a strong system of pre-service teacher education could be coupled with a stronger and more legitimate system of in-service professional development to better train teachers.

Solution 4: Stronger System of Pre-service Teacher Education

Without a robust understanding of how and when to use particular pedagogies, teachers will not be able to adopt innovative pedagogies in their own classrooms. For example, Carpenter et al. have shown that teachers without strong pedagogical training often choose pedagogies that can actually reinforce a student’s misconceptions or misunderstandings. Understanding the harmful effects of rote memorization in developing deep student knowledge could motivate

125 International Baccalaureate Organization.
teachers to utilize other pedagogies. In fact, teachers with strong pedagogical training reported feeling “pedagogical guilt,” where if they recognize just how poorly deeper knowledge is gained by rote memorization teachers attempt to use other pedagogies. 127 Another interviewee responded that every time he planned a “bad lesson where [he] just lectured” he felt guilty because he was “betraying [his] professors” and resolved to plan better lessons in the future. 128

However, there are barriers to using pre-service teacher education, which includes all of the training teachers receive before entering a classroom, as a leverage point. Across the United States, teacher education varies widely both in format and in quality. 129 Each state sets different certification requirements for teachers, and as such education programs at universities take widely different approaches to satisfy the requirements. Some university-based teacher education programs require a full degree in the subject they will be teaching, while others offer general education degrees focusing on pedagogy; some may concentrate on learning theories, psychology and sociology, while others will focus on classroom management and creating assessments. Even courses on student learning theories might teach fundamentally different concepts rooted in different areas of cognitive science. Many states also have so-called “alternate routes of teacher certification”, where teachers can have anywhere from a year to only a few months of pedagogical training, typically offered by a non-profit foundation or distance-learning courses at a university, before being placed in the classroom. As such, there is no minimum level of pedagogical knowledge that must be met before getting into the classroom.

Mandating a minimum level of pedagogical understanding is also difficult because researchers have little understanding of the preparation that teachers need before moving into the classroom. Much of the research literature about pre-service teacher education, as Wilson has shown in her meta-analysis, is done by the university itself and advocates for their particular approach to teacher education. 130 By unifying pre-service teacher education into one coherent system, educators could ensure that all teachers have the prerequisite knowledge they need before moving into the classroom. Such teacher education programs would teach a common set of pedagogical approaches and classroom management techniques, require a uniform amount of

127 Wilson, Floden, & Ferrini-Mundy, Teacher Preparation Research: Current Knowledge, Gaps and Recommendations.
128 Teacher 2
129 Wilson, Effective STEM teacher preparation, induction and professional development, (White paper to the Board on Science Education, National Research Council, 2011).
130 Wilson, Effective STEM teacher preparation, induction and professional development.
pre-service experience in classrooms and encourage a similar content knowledge background. Though each teacher will still have different strengths and weakness when they come out of pre-service education programs, districts and school would know where to begin their induction program that bridges the gap between theoretical teacher education and the realities of teaching.

Solution 5: Develop a Robust System of In-service Teacher Professional Development

In-service professional development programs, where teachers are given additional training throughout their teaching career in the form of mentorship, regular professional development days, and other learning opportunities, also have a wide range of quality and content. Some beginning teachers are part of an induction program, where experienced teachers mentor new teachers. For more experienced teachers, professional development programs, which have wide and varying quality, are provided throughout each school year. Professional development programs are often incoherent: teachers typically move from a session with one pedagogical focus to a completely disconnected training for a new technology. Teachers often have to make decisions about what training to attend without knowing the quality of the sessions or how the topics covered will fit into their existing training. According to Wilson, Rozelle and Mikeksa, this leads to little real training because the concepts taught during in-service professional development are “one-shot” and are not reinforced over time. Within this “patchy” system of professional development, the university professors, education resource developers, and other professional development providers have little idea about what teachers already know and can’t target the content of their session effectively. The lack of understanding of what teachers need to learn and how teachers learn leads the American professional development system to fall short of providing quality training.

Compounding this difficulty is the perceived lack of efficacy in professional development. One principal argued that “teachers experience pedagogical lock-in, and no amount of professional development is going to change them. We just have to wait for them to retire.” All teachers interviewed stated that they rarely changed their practices or improved their content knowledge because of professional development. Most teachers only attended professional

131 Ibid.
133 Wilson, Effective STEM teacher preparation, induction and professional development.
134 Principal 5
development because they were required to attend. One teacher reported that most professional development was not very compelling or well presented. Another teacher reported that in a typical professional development session, the new pedagogical method presented “has little evidence to support me teaching my class in that way” because of the lack of rigorous testing of many professional development lessons. Without this legitimacy, most professional development does little for teachers. Further, one teacher identified that the lack of engagement among other teachers at the professional development session takes away from the experience for other teachers, effectively “dooming most [professional development] right from the start.”

Most teachers use information from informal networks to choose their pedagogies instead of relying on professional development. Gibbons has shown that teachers get advice from their peers who are also trying new pedagogies, and trust that advice more than from any other source. These advice networks can come locally from within their school, typically a subset of their school, or shared more broadly over the internet. Interviewed teachers granted legitimacy to these informal networks, typically thinking them to “more reliably give [teachers] advice on what actually works in the classroom” than any professional development provider. One teacher reported “there is a group of teachers at my school known as the after-school club. We all stay after work and compare our experiences with one another and try to make improvements the next day... It might be more work, but we are improving outcomes for the students.” One teacher relied upon a virtual peer group, saying that “Twitter serves as my informal peer group. We get in discussions, send each other blog posts and participate in a rich online community.”

Therefore, it is important to have a robust system of in-service teacher education that also has legitimacy with teachers. The Japanese case developed later in this chapter will present an example of how such a system of professional development can be established by empowering teachers.

135 Teacher 8
136 Teacher 7
138 Teacher 6
139 Teacher 2
140 Teacher 4
5.1.3 Teacher Content Knowledge

Without strong content knowledge, teachers will still be unable to explore new pedagogies because they may not be comfortable in uncontrolled classroom settings. Teachers also spend preparation time learning the content and relying on pre-made materials rather than using new or more effective pedagogical methods. To improve this, teachers need access to better STEM training at an undergraduate level and a set of quality pre-made resources to bring experts into the classroom.

Solution 6: More Effective STEM Training in Undergraduate Education

While there is not a clear link between teacher content knowledge and student achievement, there is agreement in the literature that deeper content knowledge leads to a deeper understanding of scientific practices and better scientific pedagogies. As researchers have found, teachers lacking deep content knowledge will have difficulty using problem-based learning or other constructivist techniques. Graber, Tirosh and Glover have found that prospective teachers of science and math typically have fundamental misconceptions of practices in these fields; they can demonstrate and solve particular problems but have difficulty with more fundamental explanations of how and why phenomenon or algorithms behave in particular ways.

To better develop teacher content knowledge, classes at universities must be improved. Undergraduate classes traditionally do little to correct fundamental misunderstandings and do not develop good scientific practices in future teachers. Based primarily on lecture and rote memorization, introductory university courses harm future teachers by reinforcing incorrect beliefs about connections between topics in their fields and scientific practices. Without a

141 Wilson, Floden, & Ferrini-Mundy, Teacher Preparation Research: Current Knowledge, Gaps and Recommendations.
144 Wilson, Floden, & Ferrini-Mundy, Teacher Preparation Research: Current Knowledge, Gaps and Recommendations.
working understanding of the field, it is difficult for future teachers to translate science to their students or inspire passionate exploration of novel topics. University professors, on the other hand, are experts in their field, and do not face any content knowledge barriers in changing their pedagogies. As such, university professors should improve their pedagogical method to ensure future teachers receive a robust undergraduate education that they can pass onto students.

Solution 7: Bring Expert Materials and Knowledge into the Classroom

If it is untenable to train teachers as thoroughly as necessary to facilitate new pedagogies, then outside expertise should be brought into classrooms with new materials developed by content experts. Recall that one teacher commented that, without in-depth content knowledge, teachers had to spend valuable preparatory time on learning the content themselves instead of planning the lessons. He continued by saying "when the content is difficult for me, all of my time gets spent making sure I can answer direct questions about the material. In turn, I often end up relying on pre-made materials. If these materials facilitated project-based learning, I would do that...but they are usually lecture-based materials." Such materials would ideally provide a full lesson, where teachers can simply use a pre-made in-depth lesson, thus saving themselves time.

The MIT BLOSSOMS project would be an ideal example. BLOSSOMS is an online repository of blended learning videos where students watch a short video from a content expert in the classroom and are then guided through an activity by their teacher (for which the resources are available online) before watching another short video segment and repeating the cycle. Such materials would supplement teacher knowledge and help them engage in in-depth investigations.

5.1.4 Culture and Competition

School culture is notoriously hard to change, even with concerted effort. However, once new pedagogies are adopted, this same culture will continually support their use in the classroom. Rather than completely dismantling the reinforcing effects of school culture, if that is even possible, policies should be used to slowly change school culture and then take advantage

---

147 Teacher 5
148 Available at blossoms.mit.edu
149 Louis, Marks and Kruse, "Teachers' Professional Community in Restructuring Schools."
of the reinforcing effect once new pedagogies are implemented. Therefore, to change the culture in schools, we have to ingrain a particular culture in student teachers before they move into schools.

**Solution 8: Use Pre-service Clinical Training Programs to Ingrain a Particular Culture into Teachers**

As a part of pre-service teacher education, clinical training programs, where student teachers spend time receiving on-the-job training in classrooms, have been particularly effective at preparing teachers for the realities of every day teaching. Again, these experiences range widely, with some programs having a year of in-classroom teaching while others can have as little as 4 weeks in the classroom. However, Wilson, Floden and Ferrini-Mundy have found that most clinical programs often are poorly designed, with most clinical experiences disconnected from the real classroom experience and students receiving little mentorship from their supervising teachers. Without this mentorship, student teachers often end up focusing their clinical training on learning the procedural aspects of teaching such as discipline and workbook usage, rather than experimenting with new pedagogies. If these pre-service teachers were instead given the opportunity to see the value of new pedagogies firsthand, they would bring that knowledge with them into their first school and slowly influence the culture of schools to value the use of pedagogies other than rote memorization.

**5.2 International Use of Leverage Points to Affect Change**

The United States is not alone in having to tackle these complex problems. We can find valuable information on the effectiveness of policy changes at various leverage points by examining three international cases: Japan, Singapore and Finland. The Japanese case demonstrates how difficult changing pedagogical practice can be without changing the

---

150 Wilson, Floden, & Ferrini-Mundy, *Teacher Preparation Research: Current Knowledge, Gaps and Recommendations*.


accompanying assessments or the emphasis placed on test results by the culture. The Singaporean case demonstrates how, with strong teacher training, teacher autonomy, and a school culture that resists rote memorization, new pedagogies can be adopted alongside standardized testing, albeit with major stress on teachers. The Finnish case demonstrates how strong teacher training and teacher autonomy without standardized tests lead to new pedagogical practices that can become produce high-quality results.

5.2.1 The Japanese Experience

Japan has a centralized education system administered by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT). Compulsory education in Japan requires nine years, including six years of elementary school and three years of junior high school. Although high school is not compulsory, approximately 96% of Japanese students complete high school (with approximately 2% dropping out every year). To enter high school, students take an entrance examination at the end of the third year of junior high school.

In 2000, Japan embarked upon a number of reforms designed to increase student achievement and encourage teachers to adopt new pedagogies. However, in the last decade, Japan has not seen major change in pedagogical method as a result of the new curricula. As test scores began to fall across the country, MEXT moved away from its relaxed policy and re-adopted the old curriculum in 2011.

*Attempted Leverage Point: Reducing Pressures on Teachers to Teach to the Test*

This reform reduced the number of concepts appearing in the national curriculum by 30% as an attempt for a more “relaxed education.” As a response to widespread concerns over “cramming” for high-stakes exams, which were regarded as the root cause of school violence, bullying, drop-outs and suicides, this reform relieved some of the pressures on students by lowering the number of concepts on the test. Along with relieving student pressures, the reform also lightened the teacher workload significantly. As a result, teachers were encouraged by

---

153 Benesse, "New course goals to be decided for each school with understanding the object of the new course of study," *VIEW* 21 (2012) 4-9.

MEXT to develop students' ability to learn and think independently, not just cramming for exams.\textsuperscript{155} Many education experts interpreted this as a call for more constructivist pedagogies such as problem-based learning. However, despite the opportunity from a lightened workload to improve their practice, teachers did not widely adopt constructivist pedagogies throughout the country.

\textit{Attempted Leverage Point: Strong System of Pre-service and In-service Teacher Training}

Japan has a well prepared teaching workforce that should have had both the content and pedagogical knowledge to successfully implement the shift towards using new pedagogies.\textsuperscript{156} Under the current system, pre-service and in-service teacher training in Japan is multi-dimensional, continuous, and systematic at the national, prefectural, and more local levels. Among these mandatory training programs, there is a formal mentorship for newly appointed teachers for one year and a robust in-service professional development program for teachers with less than 10 years of experience. Under-qualified teachers take special training programs to bring their pedagogical and content knowledge up to a minimum level.\textsuperscript{157} One major element of this strong in-service professional development is the practice of Lesson Study. Lesson Study is the collaborative planning of a high quality lesson by a small group of teachers, usually four or five. Though the byproduct is a high quality lesson that can be shared by others, as implied by the name itself, the true value for teachers comes through the reflective and self-critical process of developing these lessons. Through an iterative process, individual teachers take on the role of reflective professional researcher and can observe their own actions and the actions of others more objectively.\textsuperscript{158,159,160,161} Such methods have been attempted in the United States, but typically teachers feel overwhelmed with such in-depth and continuous professional development on top of their already large workload. In Japan, however, the teacher workload

\begin{itemize}
\item \textsuperscript{155} Ibid.
\item \textsuperscript{156} Ingersoll, “The Qualifications of the Teaching Force in Japan,” In Robert Ingersoll et al eds. \textit{A Comparative Study of Teacher Preparation and Qualifications in Six Nations}, (CPRE Policy Briefs, 2007), 47.
\item \textsuperscript{157} Ibid.
\item \textsuperscript{160} Lewis, Perry, and Hurd. “A Deeper Look at Lesson Study.” \textit{Educational Leadership} 61(2004), 18-22.
\end{itemize}
does not appear to be a barrier to pursuing Lesson Study because of the relatively few hours teachers spend in front of the classroom each day, at 500 hours teaching over the course of a year in comparison to the United States' average of 1051 hours. With this method of professional development, teachers are more prepared to adopt new pedagogical approaches in the classroom.

**Barriers to Improving Pedagogy that Persisted Despite the Reform**

When MEXT lowered the institutional pressures to teach to the test, as seen in Figure 12, they expected teachers to freely experiment to explore higher learning skills. However, Japan has long prioritized the outcomes of entrance examinations to both high schools and colleges, and this culture has persisted despite the various reforms attempted in the last ten years. As such, teachers face intense pressure from students and parents to teach to the test, to a much greater extent than in the United States. This pressure is compounded by a teaching culture that emphasizes acquiescing to this pressure and “teaching to the test.” Instead of taking advantage of the reform and the additional time to explore topics more deeply, teachers were still teaching to the test, just covering less content.

Figure 9 shows the Japanese attempts at reform. Green variables, *pedagogical knowledge* and *emphasis placed on test results by administrators*, denote leverage points that might have led to the adoption of new pedagogies. Red variables, *acceptance in school culture of teaching to the test*, *continued reliance on the student's performance revealed through assessment* and *pressure on teacher to improve performance* from parents and students, are factors that stood in the way of the adoption of new pedagogical methods. This Japanese experience with reform illustrates how difficult changing pedagogical practice can be without changing the accompanying assessments or the culture in schools emphasizing those test results.

---

164 Ibid.
5.2.2 The Singaporean Experience

Although the Singaporean education system currently produces high achievement on international comparisons, they have achieved large improvements in a short time frame. Singapore only implemented its compulsory education system in 2003 and compulsory education is only six years long. After six years of primary education, there are 4 tracks of secondary education lasting 4 to 5 years, followed by 2 to 3 years of pre-university education at junior college, polytechnic or the Institute for Technical Education. Students take national examinations after each stage of school. Traditionally, the educational emphasis had been on rote memorization, rather than ability to think critically.

To improve their education system, the Ministry of Education (MOE) introduced a greater focus on critical thinking and other "life-long" skills, rather than simply learning to excel on exams. This process began with the 1997 reform, "Thinking Schools, Learning Nation"

---

followed by the 2004 “Teach Less, Learn More” policy. It has had resounding effects on their education system and improved their international rankings on measures such as the PISA exam.

**Attempted Leverage Point: Reducing Pressures on Teachers by Introducing Flexibility**

Over the last ten years, the Ministry of Education has taken numerous steps to provide more flexibility and choice for students by transforming the structure of their educational system. They have moved away from the centralized top-down system to give more autonomy to local schools and to develop new forms of accountability; each school now sets its own goals and annually assesses its own progress. Without a strong system of national accountability, there is less institutional emphasis placed on the results of standardized tests. This has eased the outside pressures on teachers and allowed them more freedom to experiment with new pedagogies.

**Attempted Leverage Point: Strong System of Pre-service and In-service Teacher Training**

MOE has been trying to build up a strong qualified teaching force through a process of careful and detailed planning, aggressive teacher recruitment, comprehensive training and effective teacher retention, both quality and quantity. All teachers are hired by MOE from the top one-third of each class of college graduates. All teachers and trainee teachers receive pre- and in-service training at the National Institute of Education (NIE) to ensure adequate content and pedagogical knowledge in the classroom. NIE works symbiotically with MOE to give advice on hiring of teachers and jobs are guaranteed to those who complete the NIE pre-service training. The MOE has also provided resources and support for schools, including Professional Learning Communities (PLCs) in schools and Centers of Excellence to “facilitate sharing of good teaching practices among teachers and schools.”

**Barriers to Improving Pedagogy that Persisted Despite the Reform**

This reform has generally been considered a success. Though Singapore has retained high stakes standardized tests, studies show that the classroom experience has begun to move away

---

169 Ibid.
from rote memorization. However, it is also reported that teachers are under high levels of stress as they try to educate students both with deep pedagogical practices and for the standardized tests. Pressure from standardized tests remains intense, with students and teachers still valuing the results of test scores more than true student knowledge. According to Zhao, the sociocultural practice of competition is wearing upon teachers, and it is unclear how much longer they will continue to practice effective pedagogical methods when there is still intense pressure from students and parents to improve test scores. Such strong pressure could undo all of the progress Singapore has made if teachers begin to compromise their pedagogies to appease students or parents.

Figure 10 shows the Singaporean attempts at reform. Green variables, pedagogical knowledge, teacher content knowledge, acceptance in school culture of teaching to the test and emphasis placed on test results by administrators, denote leverage points that might have led to the adoption of new pedagogies. Red variables, continued reliance on the student's performance revealed through assessment and pressure on teachers to improve performance from parents and students, are factors that might have stood in the way of the adoption of new pedagogical methods. This Singaporean case demonstrates how, with strong teacher training, teacher autonomy and developing a school culture that resists rote memorization, new pedagogies can be adopted alongside standardized testing, albeit with major stress on teachers.

\[^{170}\text{OECD, "Singapore: Rapid Improvement Followed by Strong Performance."}\]
\[^{171}\text{Zhao, Handbook of Asian Education: A Cultural Perspective, (Routledge, 2011).}\]
\[^{172}\text{Okubo, Hird, Uchino and Larson, "An Examination of Learning Standards in a Comparative Perspective: U.S., Japan, Singapore, and Finland."}\]
\[^{173}\text{Ibid.}\]
5.2.3 The Finnish Experience

Compulsory education in Finland is 9 years, from the age of 7 until age 16. Overall, 98% of children attend preschool, but early childhood education is not mandatory. The completion rate of compulsory education is 98%. If students score multiple failing grades during that time, they may have to repeat the year (with pupil and parental consent). All school-related expenses, such as school healthcare, lunch, books and materials, and school trips, are free during the first nine years. At the age of 16, students choose whether to undergo occupational training to develop vocational competence or to enter an academic upper school for university and postgraduate professional degrees. Upper secondary schools are 3 to 4 years long. Finnish class sizes are small, with approximately 20 students per class, and they further divide students into smaller working groups for long-term projects. Students stay with the same teachers for several years, and there is no ability-based tracking.

---

175 OECD, “Finland: Slow and Steady Reform for Consistently High Results.”
From 1970 to 1985, Finland centralized education standards and adopted their first national standards. Since 1990, however, Finland has focused on decentralizing power, giving teachers control over their teaching methods by only having a framework curriculum.\textsuperscript{176} Finland is of particular note because of its results on international comparisons, routinely scoring at the top of the PISA and TIMSS evaluations.

\textit{Attempted Leverage Point: Removing Standardized Testing}

In Finland, there is only one standardized test at the age of 16. There are no state mandated tests every year, and assessments are primarily given both formally and informally by teachers. Further, students are not measured in comparison to others in the first six years.\textsuperscript{177} Instead, Finland places a focus on using subjective assessment. The classes are designed so that lessons are based more on student input (60% student input, 40% teacher input); they focus on student-centered learning and let students discover concepts rather than lecturing to teach those concepts. These constructivist lessons are based on debate, and there is very little homework. Finland clearly illustrates that the pedagogies used in classrooms can be changed, especially when assessments are not in place driving teachers to teach to the test. In short, learning is the end goal of the education system, not simply the results of a particular test.

\textit{Attempted Leverage Point: Strong System of Pre-service and In-service Teacher Training}

A large part of Finland’s success seems to come from their strong pre-service professional development program. A Master’s degree is required for all teachers, requiring both in-depth content knowledge and general understanding of education principles and pedagogies, and teachers are typically from the top 10\% of their graduating class. In addition, they have to demonstrate their content knowledge and pedagogical skills through writings and discussions in the examination to be hired as teachers. They used higher standards to influence teachers of all abilities to get on board. These high standards raised the interest of future students in higher education.

\textsuperscript{176} OECD, “Finland: Slow and Steady Reform for Consistently High Results.”
\textsuperscript{177} Hancock, “Why are Finland’s School Successful?: The Country’s Achievements in Education have Other Nations Doing Their Homework.” \textit{Smithsonian Magazine} (September 2011).
education, eventually depressing the wages of a master’s degree enough to allow Finland to require a Master’s degree.\textsuperscript{178}

\textit{Attempted Leverage Point: Developing Culture to Support Teachers}

In Finland the national curriculum is set as broad guidelines for teachers, and the public and policymakers place a lot of trust on teachers as professionals. Teachers have a culture of equity and equality, and they believe in a high degree of personal responsibility and individuality, emphasizing caring and cooperation with others rather than competition. There is also more focus on “learning to learn” (or critical thinking) than on learning the subject matter. This culture has been instilled in part because of an extensive system of clinical pre-service training. As a result, the reform has created a culture inside of schools where teachers are “researchers” and experiment with developing and perfecting their own pedagogies.

Figure 11 shows the Finnish attempts at reform. Green variables, increasing pedagogical knowledge, and teacher content knowledge, lowering acceptance in school culture of teaching to the test and emphasis placed on test results by administrators, removing reliance on student’s performance revealed through assessment and removing pressure on teacher to improve performance from parents and students, are factors that led to the adoption of new pedagogies. This Finnish case demonstrates how strong teacher training and teacher autonomy without standardized tests leads to new pedagogical practices that can become ingrained in a system to produce high-quality results.\textsuperscript{179}

\textsuperscript{178} OECD, “Finland: Slow and Steady Reform for Consistently High Results.”
\textsuperscript{179} Okubo, Hird, Uchino and Larson, “An Examination of Learning Standards in a Comparative Perspective: U.S., Japan, Singapore, and Finland.”

74
Figure 11: Finnish Attempts with Reforms to use new Pedagogies
6. Policy Recommendations

Now that we have identified individual leverage points and examined how these leverage points have been used both effectively and ineffectively, we can develop policies to improve STEM education in the United States. However, such policy recommendations will have to influence a system that is already in the midst of changing policies. In response to stagnating standardized tests scores, policymakers have developed new state standards of performance that each student should be able to meet at each grade level. In Mathematics, the Common Core State Standards have been adopted by 45 states; in science, the Next Generation Science Standards (NGSS), released in April 2013, are currently being considered for adoption by some states.

By examining these new standards in the context of our causal model we can see where policy shifts will have impacts on certain leverage points and how those impacts will resonate throughout the rest of the model. Once I have developed predictions of how the NGSS will affect the system, I will develop more generalized policy recommendations that will work in tandem with these new standards to improve educational outcomes for all students in the STEM fields.

6.1 Impact of the Next Generation Science Standards

The goal of the NGSS is to circumvent the ability for teachers to use lecture-based pedagogies to improve test scores. The NGSS are made up of not only the subject matter and topics that are typically contained in a curriculum, but also the set of scientific practices (the set of skills that scientists and engineers use) and crosscutting concepts (such as cause and effect, systems and system models, etc.) that students should understand. In theory, the NGSS take a step towards aligning expected student performance with the development of deep student knowledge; however, as developed in the causal model, the standardized tests developed from the standards will be the factor that drives pedagogical selection. Therefore, no matter how much these standards theoretically incentivize new pedagogies, the fidelity of the standardized tests to the standards will determine actual teacher behavior. The NGSS are an attempt to follow Solution 2 as developed in Chapter 5 and improve the quality of standardized tests, which will, in theory, more closely align standardized tests with student knowledge.

---

Developing new assessments that can measure the types of performance standards in the NGSS is not going to be a simple task. One interviewee was particularly adamant that “the new standards are going to be incredibly difficult to assess. Just look at PARCC and SBAC [the assessment consortia currently developing new assessments for the Common Core State Standards in Math and Reading] who are having such large difficulties developing in-depth tests for what, I think to be, easier and more straightforward standards [in Math and Reading].” Many in education agree; in a recent survey of education policy experts, more than 2/3 thought that neither group was on the right track to producing quality assessments for the Math or Reading Common Core State Standards. If these test builders, which are supported by federal funds, are having difficulties developing such rigorous standardized tests for comparatively easier content to assess, assessment designers for the NGSS are going to have even more difficulties.

There are, however, many other leverage points in the system that are not being addressed by these new standards. As seen in the three international comparisons in Chapter 5, reforms that only try to leverage one or two changes are often not effective. The range in quality of pre-service teacher training will continue to impede the ability of some teachers to lead the deep investigations demanded by the NGSS. Further, little is being done to change the culture of schools that are already using rote memorization to improve test scores.

Figure 12 illustrates the set of leverage points most relevant to the implementation of the NGSS. The yellow variable in this diagram represents a leverage point which could either support new pedagogies or reinforce the usage of poor pedagogies. Here, a student’s performance revealed through assessment depends heavily on the quality of assessments developed. If test designers are able to sufficiently improve assessments so that they would impede efforts to use lecture-based pedagogies, unintended consequences of the pressure on teachers to improve performance would be eliminated. However, if the quality of assessments only improves marginally, there will still be strong incentive for teachers to continue to use rote memorization to improve these test scores. Added to this situation are the red variables, representing barriers to change that are persisting. Teacher training remains unchanged and thus pedagogical knowledge and teacher content knowledge will remain as barriers. The acceptance

---

175 Administrator 1
176 Whiteboard Advisors, “Tracking Measures, Common Core Materials, and Other Timely Topics in Education.”
in school culture of teaching to the test remains unchanged and policymakers continue to value the objectivity of measures, both reinforcing the use of lecture-based pedagogies in schools. Even with high-quality assessments, which is unlikely, the numerous factors reinforcing the reliance on rote memorization will overpower any major changes brought in by the Next Generation Science Standards.

Figure 12: Unchanged Leverage Points under NGSS that will impede changes in pedagogy

6.2 Generalized Policy Recommendations

Through a further examination of successes and failures of international experiences with pedagogical reform, I have developed various policy recommendations that will complement the adoption of the Next Generation Science Standards and interrupt the causal factors pushing teachers to adopt rote memorization in their classroom. Broadly, new policies must interrupt the shortcut that teachers may take using lecture-based pedagogies, such as rote memorization, to improve test scores. Teacher training programs must also be strengthened to provide support for their use of constructivist pedagogies.
6.2.1 Remove Standardized Tests and Develop a Legitimating System for Subjective Tests

Pouring money into the development of objective assessments for the Next Generation Science Standards is ineffective when there are less expensive and well tested options. Instead, it would be more effective to adopt two other solutions in tandem: removing standardized tests and developing a system of many multiple measures, including subjective tests graded anonymously by teachers around the state. As the case of reform in Finland shows, it is entirely possible to have a high-performing education system without putting nearly as much emphasis on large-scale objective assessments. Once the objective standardized tests are removed, a system of subjective tests should be put in place similar to the International Baccalaureate tests as detailed in Chapter 4. These exams would be mainly subjective and ask in-depth questions to effectively assess student knowledge. To provide accountability and legitimacy, these tests should be graded by other teachers around the state. As such, there would still be pressures to improve student outcomes, but these tests would no longer be able to be “gamed” with lecture-based pedagogies.

These subjective assessments could also serve as a main component of the portfolio approach to student assessment as developed in Chapter 5. Again, such a portfolio assessment filled with many end-of-semester projects, well-written essays, and teacher evaluations of deeper skills such as critical thinking and persistence could be used to provide a detailed account of student knowledge to students, parents, and college admissions offices.

6.2.2 Implement Lesson Study as Improved Pedagogical Training

Poor teacher opinion of in-service professional development holds back the effectiveness of such training. This is a vicious cycle, where low quality programs discourage teachers, which hinders what would otherwise be effective professional development. To break out of this cycle, a strong system of professional development is needed, such as Lesson Study, as seen in the Japanese case. Teachers focus on creating high-quality lessons in groups of 3-5 over a period of a few months. The teachers not only see the direct results of their professional development in the lesson they plan and test in their classrooms, but develop more introspective habits as well. This collaboration between teachers can also spill over into other aspects of teaching, where teacher teams can build upon each other’s expertise to co-teach lessons that use new pedagogies.
A valuable outcome of the Lesson Study process is the large number of extremely high quality lessons, which could be shared with other teachers around the country, relieving the workload of each.

6.2.3 Improve Pedagogical Practice at Universities to Improve Future Teacher Knowledge

Lecture-based teaching at universities is holding future teachers back. With the predominant use of lecturing in introductory university classes, the most advanced STEM classes that most future teachers will take typically reinforce incorrect beliefs about connections between topics in their fields and scientific practices. Instead, universities need to provide more authentic opportunities for future teachers to develop deeper content knowledge. Undergraduate professors should be encouraged to adopt innovative pedagogies, which would train teachers more deeply and serve as a model to be imitated by future teachers.

Another of the most effective methods of training undergraduates has been undergraduate research internships. At these internships, future teachers can apply their in-classroom knowledge and begin to understand the methods scientists and engineers use firsthand. Implementing such programs for all STEM teachers would undoubtedly improve their content knowledge and allow them to better understand the scientific method.177

6.2.4 Change the Culture in Schools through Pre-service Clinical Experiences

Even if all assessment practices were changed to allow teachers the freedom to experiment, teachers will not be prepared to change their practice without pre-service clinical training. The Singaporean case also demonstrates the power that such clinical training programs can have in developing a school culture that resists strong pressures to improve test scores. To break the secondary reinforcing loops, where culture and competition reinforce rote memorization, strong pre-service teacher training will have to instill well-developed practices and a culture of pedagogical experimentation. This well-developed self-reflective practice among teachers will decrease the reliance on rote memorization, in turn reinforcing a school culture of pedagogical experimentation. Federal and state policies mandating standardized testing would be

opposed by a strong culture opposing teaching to the test, which would lead to even greater achievement gains. The most likely opportunity to intervene in this self-reinforcing loop is in formalizing the haphazard network of pre-service teacher education that exists across the country.

6.3 Future Work

The causal interactions developed and applied in this work could be used to support further modeling efforts in education. Though this work has stayed away from developing a full Systems Dynamics model because of the difficulties in using accurate measures, it does not preclude such a model from being developed. The largest difficulty with the development of such a model is using a dataset that is reliable. Most large datasets of testing information are aggregated at a school level, making them difficult to use in a model describing individual teacher actions. Another approach may be to use a small number of case studies, probably schools that have had sharp increases in test scores when a certain policy was implemented compared to schools that have implemented the same policy without any effect. These small N case studies typically attribute the cause of such implementation failures to a single factor that has little value from a system-wide point of view. However, with enough in-depth access to various classrooms, combined with a thoughtful research design, there could be enough data gathered to develop a full working model that can inform policy decisions.

At a system-wide level, the adoption and implementation of the Next Generation Science Standards provides a natural experiment for the validation of this causal model. Twenty-six states were a part of the development process; many of which will adopt the standards. However, each state will develop their own policies to support the implementation, which will pull on different leverage points identified in this model. Using a case study methodology, states with different approaches to the critical leverage points could be identified and compared to better describe how state policy actions filter down to the classroom context.

---

178 Elmore, “Structural Change and Educational Practice.”
6.4 Conclusions

This research used a series of interviews with educators at all levels in the system to identify the guiding factors in the adoption of, or lack thereof, new pedagogical methods. A causal diagram was developed from these interviews, leading to the conclusion that great pressure placed on teachers, both institutionally from schools and from parents and students, causes them to adopt the pedagogy most well suited to improve test scores, regardless of the long-term impacts of this pedagogy on student knowledge. Specific leverage points in this model were identified, and international attempts at using these leverage points were examined to develop policy recommendations. With insights from this research, education policymakers will better be able to understand how future policy decisions will impact the system, as well as the critical factors that will continue to hold back progress towards providing a better future for all students.
Bibliography


Benesse, 「新課程の趣旨を踏まえ 自校に必要な重点項目を定める」(Deciding important agenda for each school with understanding the object of the new course of study). VIEW 21 (2012), 4-9.


Gerwertz, C., "Publisher Addresses Common-Core ‘Alignment Issue.’" *Education Week*. Published April 11, 2012.


Hancock, L., "Why are Finland’s School Successful?: The Country’s Achievements in Education have Other Nations Doing Their Homework," *Smithsonian Magazine*, September 2011.


MEXT, 総合的な学習の時間専門部会における審議状況について (Deliberations in the Special Committee on the status of “integrated learning” time), 2005.


National Commission on Teaching and America’s Future, Unraveling the “Teacher Shortage” problem: Teacher Retention is the Key, 2002.


National Research Council, Assessment in support of instruction and learning: Bridging the gap between large-scale and classroom assessment, Workshop report, Committee on Assessment in Support of Instruction and Learning, Board on Testing and Assessment, Committee on Science Education K-12, Mathematical Sciences Education Board, Center for...


Ng, J. and S. Sreedharan, Teach less, learn more - have we achieved it? Outcome is a mixed bag as some children are still being pressured to do well in exams, TODAY, August 24, 2012.


OECD, Programs for International Student Assessment (PISA) 2009 Results, 2009.

Okubo, “International Comparison of Education in Asia and Finland – Understanding Math and Science Education in Japan, South Korea, China, Taiwan, Singapore and Finland,” Unpublished document submitted to Dr. Richard Larson as a part of Towards Intelligent Societies: What Motivates Students to Study Science and Math? How Do We Provide for Flexible Learning Pathways? 2012.


Partnership for the Assessment on Readiness for College and Careers, Summative Assessments at the High School Level.


Soares, SAT Wars: The Case of Test-Optional College Admissions, (Teachers College Press, 2011).


Thayer, T., *Finland's educational policy environment: The role of strategic ambiguity in policy communication*, PhD Presentation, University of Iceland, 2012.


Ujifusa, Andrew, “Pressure Mounts in Some States Against the Common Core,” Education Week, February 6, 2013.


Whiteboard Advisors, “Tracking Measures, Common Core Materials, and Other Timely Topics in Education,” April 2013.


Wilson, S., Effective STEM Teacher Preparation, Induction and Professional Development. Commissioned Paper for the National Research Council’s Board on Science Education “Highly Successful Schools or Programs for K-12 STEM Education: A Workshop,” 2011.


Appendix A: Survey Instrument

<table>
<thead>
<tr>
<th></th>
<th>Strongly Agree</th>
<th>Moderately Agree</th>
<th>Slightly Agree</th>
<th>Neither Agree nor Disagree</th>
<th>Slightly Disagree</th>
<th>Moderately Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>As a pedagogical model, blackboard chalk and talk does not work well with today's students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Getting students to co-teach among themselves increases interest, engagement and learning outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“One size fits all” lock step teaching for all students in a class slows and discourages advanced students and can make slower students fall further behind</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students should be separated into different streams by ability and taught in classrooms of similar ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One poor math or science teacher in the K-12 chain of classes can derail an otherwise promising student career in STEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One great math or science teacher in the K-12 chain of classes can motivate a student to pursue a career in STEM regardless of the teachers they have in the next grade levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers need more training in innovative and productive ways of preparing for class and teaching</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers are well trained in innovative and productive but the current incentive structure does not encourage them to use newer methods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers require substantial outside assistance from scientists and engineers in creating the most effective STEM learning environments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Those who would be gifted STEM teachers are difficult to recruit into the teaching profession because of low pay in comparison to traditional STEM jobs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEM teachers should be paid more than teachers of other subjects to acknowledge difficulties in recruitment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students’ need integrative critical thinking skills, across traditional silo boundaries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The current organization of STEM subjects into Physics, Chemistry, Biology and Mathematics is the most efficient way to teach these subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardized tests evaluate algorithmic steps or facts, not critical thinking skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardized tests provide important information about student performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intelligent use of technology in education can attract more students to STEM careers and speed the learning process</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning by game playing is a way of engaging young people in math and science principles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small schools generate better results than larger schools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The “gifted” students should be the focus of STEM education in order to challenge them and motivate them to continue improving</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The focus on STEM K-12 education should be on those not intending to go to elite colleges and universities but on the middle of our emerging workforce that needs more numeracy, more familiarity with the scientific method</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent attitudes about STEM are the most critical factor in motivating their children to pursue a STEM degree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B: Administrators and Principals Interview Protocol

PEDAGOGY

* Are particular pedagogies more effective at teaching STEM subjects than others?
* How can technology best be used in education?
* Is technology critical for STEM education? How does it help education?
Is there a way that social media could be used in STEM classes? If so, how?
* Do you think STEM education can help students become more analytical, critical, and creative?
How can STEM education contribute to developing these abilities?

TEACHERS

* What kind of teachers do we need for the successful STEM education?
* Are STEM teachers different than regular teachers? If so, how?
* How can we recruit the best students into STEM teaching?
* Should there be specific reward schemes set up to acknowledge the difficulties in recruitment in specific subjects?

STANDARDS AND TESTING

* Who is considered a success in STEM education? Are students who enroll in a university STEM degree, students who complete a STEM degree at a college or university and students who is competent and interested in science but choose to pursue a non-STEM degree seen differently?
* What are attributes of a successfully STEM student?
* Some attempts have been made to organize STEM subjects by subject, such as environmental science, rather than the traditional Physics, Chemistry, Biology and Math. Does this help STEM education? What are barriers to implementing this for all subjects?
* Is the aim of STEM education to produce students that will pursue a STEM career or simply a student prepared for college?
* If standardized testing requirements were to vanish, how would you educate differently?
* What information do standardized tests (in any form) provide you with?
*If improvements could be made, how would you change standardized tests to provide you with better information?

*Are improved test scores necessary for the success of a STEM program?

* How should the success of the program or any others be measured (if not by standardized tests)?

* How does a national set of standards help or hinder STEM education?

* How do Technology and Engineering fit into STEM Education at the school?

* How do subjects other than STEM subjects fit into the school day?

**OTHER FACTORS**

* How critical are parental opinions to the schooling experience?

* How can we change parental opinions, particularly in the STEM fields?

* How do you view connections with community groups, businesses, universities etc.?

* Are these partnerships short term to help start programs in STEM education or longer term partnerships?
Appendix C: Teacher Interview Protocol

Pedagogy: Factors that influence actions in the classroom

* Are particular pedagogies more effective at teaching STEM subjects than others? Are lecturing and rote learning necessary as pedagogies?

* How do you learn about new pedagogical techniques? How do you evaluate a new pedagogy’s effectiveness?

* How can technology best be used in the classroom? Can technologies allow or hinder your use of some pedagogies?

* Do you build opportunities for increasing student interest into your lessons? How do you balance this with delivering content?

* What factors influence your design of a lesson plan? How much is drawn from the standard curriculum? From colleagues? Other sources?

How do Standards and Testing guide teacher behavior

* How often do you use any sort of “common materials” to plan lessons?

* Do you have enough time to teach the content that is asked of you by the curriculum? If you had to teach fewer topics, how would your teaching change?

* When your district or state changes the curriculum how do you learn about the changes? What resources help you to make the regular changes in curriculum?

* How do you balance the demands of standards/curricula and end of year assessments but still retain flexibility to adapt to the needs of your students?

* If standardized testing requirements were to vanish, how would you educate differently?

* What information do standardized tests (in any form) provide you with? If improvements could be made, how would you change standardized tests to provide you with better information?

Teacher Workforce and Motivations

* What attributes or skills sets make a STEM teacher successful?
*Are the attributes or skill sets of effective STEM teachers different than teachers of other subjects? If so, how?
* Why are teachers motivated to teach?
  How can we recruit great new STEM teachers?
  Should there be specific reward schemes set up to acknowledge the difficulties in recruitment in specific subjects?
* What discourages teachers from teaching?

Professional Development: Is it Effective?
* Why do you attend professional development?
* Do you find professional development most useful when it is focuses on new content in the field, new pedagogies, how to use new resources, etc?
* Do you have the ability to choose your own professional development? What factors go into your choice?

The Goals of STEM Education
* How do you define success in your classroom? What does it mean for a student to be successful in STEM? Mastery of Content? Interest in STEM fields? STEM-literacy?
* Can STEM education help students become more analytical, critical, and creative?
* Some attempts have been made to organize STEM subjects by subject, such as environmental science, rather than the traditional Physics, Chemistry, Biology and Math. Does this help STEM education? What are barriers to implementing this for all subjects?
* How do Technology and Engineering fit into STEM Education at the school?

Other
* There are numerous visions of what makes an effective principal, but how do you think a principal should best lead a school?
* How critical are parental opinions to the schooling experience?
* How can we change parental opinions, particularly in the STEM fields?
## Appendix D: Survey Results

<table>
<thead>
<tr>
<th>Standardized tests provide important information about student performance</th>
<th>Strongly Agree</th>
<th>Moderately Agree</th>
<th>Slightly Agree</th>
<th>Neither Agree nor Disagree</th>
<th>Slightly Disagree</th>
<th>Moderately Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 0%</td>
<td>2 13%</td>
<td>5 31%</td>
<td>4 25%</td>
<td>5 31%</td>
<td>0 0%</td>
<td>0 0%</td>
</tr>
<tr>
<td>Standardized tests provide important information about student performance</td>
<td>9 56%</td>
<td>5 31%</td>
<td>2 13%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
</tr>
<tr>
<td>Learning by game playing is a way of engaging young people in math and science principles</td>
<td>4 25%</td>
<td>10 63%</td>
<td>2 13%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
</tr>
<tr>
<td>Small schools generate better results than larger schools</td>
<td>2 13%</td>
<td>0 0%</td>
<td>3 19%</td>
<td>8 50%</td>
<td>3 19%</td>
<td>0 0%</td>
<td>0 0%</td>
</tr>
<tr>
<td>The &quot;gifted&quot; students should be the focus of STEM education in order to challenge them and motivate them to continue improving</td>
<td>0 0%</td>
<td>0 0%</td>
<td>2 13%</td>
<td>2 13%</td>
<td>2 13%</td>
<td>0 0%</td>
<td>10 63%</td>
</tr>
<tr>
<td>The focus on STEM K-12 education should be on those not intending to go to elite colleges and universities but on the middle of our emerging workforce that needs more numeracy, more familiarity with the scientific method</td>
<td>2 13%</td>
<td>1 6%</td>
<td>2 13%</td>
<td>8 50%</td>
<td>2 13%</td>
<td>0 0%</td>
<td>1 6%</td>
</tr>
<tr>
<td>Parent attitudes about STEM are the most critical factor in motivating their children to pursue a STEM degree</td>
<td>0 0%</td>
<td>3 19%</td>
<td>9 56%</td>
<td>2 13%</td>
<td>2 13%</td>
<td>0 0%</td>
<td>0 0%</td>
</tr>
<tr>
<td></td>
<td>Strongly Agree</td>
<td>Moderately Agree</td>
<td>Slightly Agree</td>
<td>Neither Agree nor Disagree</td>
<td>Slightly Disagree</td>
<td>Moderately Disagree</td>
<td>Strongly Disagree</td>
</tr>
<tr>
<td>-----------------------------------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
<td>---------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>As a pedagogical model, blackboard chalk and talk does not work well with today's students</td>
<td>4 25%</td>
<td>4 25%</td>
<td>2 13%</td>
<td>3 19%</td>
<td>1 6%</td>
<td>1 6%</td>
<td>1 6%</td>
</tr>
<tr>
<td>Getting students to co-teach among themselves increases interest, engagement and learning outcomes</td>
<td>7 44%</td>
<td>5 31%</td>
<td>2 13%</td>
<td>2 13%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
</tr>
<tr>
<td>&quot;One size fits all&quot; lock step teaching for all students in a class slows and discourages advanced students and can make slower students fall further behind</td>
<td>11 69%</td>
<td>5 31%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
</tr>
<tr>
<td>Students should be separated into different streams by ability and taught in classrooms of similar ability</td>
<td>0 0%</td>
<td>0 0%</td>
<td>3 19%</td>
<td>2 13%</td>
<td>3 19%</td>
<td>7 44%</td>
<td>1 6%</td>
</tr>
<tr>
<td>One poor math or science teacher in the K-12 chain of classes can derail an otherwise promising student career in STEM</td>
<td>2 13%</td>
<td>3 19%</td>
<td>5 31%</td>
<td>2 13%</td>
<td>2 13%</td>
<td>2 13%</td>
<td>0 0%</td>
</tr>
<tr>
<td>One great math or science teacher in the K-12 chain of classes can motivate a student to pursue a career in STEM regardless of the teachers they have in the next grade levels</td>
<td>6 38%</td>
<td>4 25%</td>
<td>6 38%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
</tr>
<tr>
<td>Teachers need more training in innovative and productive ways of preparing for class and teaching</td>
<td>3 19%</td>
<td>5 31%</td>
<td>4 25%</td>
<td>0 0%</td>
<td>4 25%</td>
<td>0 0%</td>
<td>0 0%</td>
</tr>
<tr>
<td>Teachers are well trained in innovative and productive pedagogies but the current incentive structure does not encourage them to use newer methods</td>
<td>4 25%</td>
<td>3 19%</td>
<td>2 13%</td>
<td>2 13%</td>
<td>4 25%</td>
<td>1 6%</td>
<td>0 0%</td>
</tr>
<tr>
<td>Teachers require substantial outside assistance from scientists and engineers in creating the most effective STEM learning environments</td>
<td>Strongly Agree</td>
<td>Moderately Agree</td>
<td>Slightly Agree</td>
<td>Neither Agree nor Disagree</td>
<td>Slightly Disagree</td>
<td>Moderately Disagree</td>
<td>Strongly Disagree</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>0 0%</td>
<td>0 0%</td>
<td>1 6%</td>
<td>0 0%</td>
<td>13 81%</td>
<td>2 13%</td>
<td>0 0%</td>
<td></td>
</tr>
</tbody>
</table>

| Those who would be gifted STEM teachers are difficult to recruit into the teaching profession because of low pay in comparison to traditional STEM jobs | | | | | | | |
|---|---|---|---|---|---|---|
| 4 25% | 3 19% | 4 25% | 0 0% | 4 25% | 1 6% | 0 0% |

| STEM teachers should be paid more than teachers of other subjects to acknowledge difficulties in recruitment | | | | | | | |
|---|---|---|---|---|---|---|
| 4 25% | 2 13% | 6 38% | 2 13% | 2 13% | 0 0% | 0 0% |

| Students' need integrative critical thinking skills, across traditional silo boundaries | | | | | | | |
|---|---|---|---|---|---|---|
| 13 81% | 3 19% | 0 0% | 0 0% | 0 0% | 0 0% | 0 0% |

| The current organization of STEM subjects into Physics, Chemistry, Biology and Mathematics is the most efficient way to teach these subjects | | | | | | | |
|---|---|---|---|---|---|---|
| 0 0% | 0 0% | 2 13% | 1 6% | 6 38% | 4 25% | 3 19% |

| Standardized tests evaluate algorithmic steps or facts, not critical thinking skills | | | | | | | |
|---|---|---|---|---|---|---|
| 6 38% | 4 25% | 3 19% | 0 0% | 3 19% | 0 0% | 0 0% |
Appendix E: Time Spent Teaching across OECD Countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of hours teaching in front of the classroom per year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chile</td>
<td>1232</td>
</tr>
<tr>
<td>United States</td>
<td>1051</td>
</tr>
<tr>
<td>Scotland</td>
<td>855</td>
</tr>
<tr>
<td>Mexico</td>
<td>843</td>
</tr>
<tr>
<td>Australia</td>
<td>797</td>
</tr>
<tr>
<td>Portugal</td>
<td>770</td>
</tr>
<tr>
<td>Netherlands</td>
<td>750</td>
</tr>
<tr>
<td>Ireland</td>
<td>735</td>
</tr>
<tr>
<td>England</td>
<td>714</td>
</tr>
<tr>
<td>Germany</td>
<td>713</td>
</tr>
<tr>
<td>OECD average</td>
<td>701</td>
</tr>
<tr>
<td>Spain</td>
<td>693</td>
</tr>
<tr>
<td>Slovak Republic</td>
<td>645</td>
</tr>
<tr>
<td>Belgium (Fl.)</td>
<td>642</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>634</td>
</tr>
<tr>
<td>Slovenia</td>
<td>633</td>
</tr>
<tr>
<td>France</td>
<td>628</td>
</tr>
<tr>
<td>Italy</td>
<td>619</td>
</tr>
<tr>
<td>Belgium (Fr.)</td>
<td>610</td>
</tr>
<tr>
<td>Korea</td>
<td>605</td>
</tr>
<tr>
<td>Hungary</td>
<td>597</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>595</td>
</tr>
<tr>
<td>Austria</td>
<td>589</td>
</tr>
<tr>
<td>Estonia</td>
<td>578</td>
</tr>
<tr>
<td>Turkey</td>
<td>567</td>
</tr>
<tr>
<td>Finland</td>
<td>550</td>
</tr>
<tr>
<td>Iceland</td>
<td>547</td>
</tr>
<tr>
<td>Israel</td>
<td>524</td>
</tr>
<tr>
<td>Norway</td>
<td>523</td>
</tr>
<tr>
<td>Japan</td>
<td>500</td>
</tr>
<tr>
<td>Poland</td>
<td>486</td>
</tr>
<tr>
<td>Greece</td>
<td>426</td>
</tr>
<tr>
<td>Denmark</td>
<td>377</td>
</tr>
</tbody>
</table>