

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Cultures and Contexts of Data-Based Decision-Making in Schools

**Permalink**

<https://escholarship.org/uc/item/9nn9z1q0>

**Author**

Ho, Jennifer Elizabeth

**Publication Date**

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Cultures and Contexts of Data-Based Decision-Making in Schools

A dissertation submitted in partial satisfaction  
of the requirements for the degree Doctor of Philosophy  
in Education

by

Jennifer E. Ho

© Copyright by

Jennifer E. Ho

2016

## ABSTRACT OF THE DISSERTATION

### Cultures and Contexts of Data-Based Decision-Making in Schools

by

Jennifer E. Ho

Doctor of Philosophy in Education

University of California, Los Angeles, 2016

Professor Christina A. Christie, Chair

“Data-based decision-making” or “evidence-based decision-making” in education are now popularized phrases to describe the systematic collection and analysis of various types of data to help improve the success of students and schools (Marsh, Pane, & Hamilton, 2006). The theory of action underlying data use activities implies that education practitioners who ground their decisions in evidence will more effectively deliver methodical improvements to teaching and learning. However, very little research has been conducted to test this hypothesis. In addition to the research community’s vague understanding of how schools – and the individuals comprising schools – interpret and implement data-based decision-making policies, it is difficult to determine whether data use practices are actually associated with improved instruction. As a result, school districts, as well as state and federal policy makers, have little understanding of how schools are actually using data, how differences in data use may affect school performance, and/or what kinds of measures could be used to indicate the effective use of data in schools.

This comparative case study of three high schools in Los Angeles Unified School District develops an illustrative understanding of how school decision-makers (i.e., teachers, principals, and district personnel) make meaning of directives to “use data for decision-making” and how the use of school-based data takes place in practical application. Drawing upon interview and observational data from principals, teachers, and district managers, it acknowledges that schools are inundated with multiple data sources and that teachers and administrators regularly rely on data use practices. The expectation that schools should more systematically, formally, and cooperatively review data to steer conversations around teaching and learning, however, implies paradigmatic shifts in the ways that data are currently understood and utilized.

Findings suggest that the effective use of school data in decision-making by school practitioners was not the product of an organized, rational process, nor one simply improved with the introduction of inputs and interventions. Rather, it suggests that culturally derived definitions of credible data, leadership, decision-making processes, accountability, organizational learning, and evaluation – and even whether data are relevant in teachers’ thinking in institutional contexts – shape stakeholder attitudes toward data use in classrooms and schools. In constant dialogue, stakeholders tacitly and explicitly negotiated what data were used in measuring school, teacher, and student performance, how data were collected and analyzed in ways that maintained credibility, who was involved in decision-making moments and in what ways, and how data could meaningfully inform programmatic student supports and instructional improvements. Data and data use processes intended to influence decision-making were, as a result, reliant on cultural, political, and subjective factors, and evolved in necessarily gradual cycles of establishment, revision, and refinement.

This dissertation of Jennifer E. Ho is approved.

Marvin C. Alkin

Kathryn M. Anderson

Bruce Fuller

John S. Rogers

Christina A. Christie, Committee Chair

University of California, Los Angeles

2016

To my mom and dad, lifetime models of commitment and dedication;  
my husband, unwavering in his selflessness, support, and enthusiasm;  
and to Tyler and Wyatt, who have shown me we can all do beyond what we believe is possible.

# Table of Contents

|                                                           |            |
|-----------------------------------------------------------|------------|
| <b>LIST OF TABLES .....</b>                               | <b>X</b>   |
| <b>LIST OF FIGURES .....</b>                              | <b>XI</b>  |
| <b>ACKNOWLEDGMENTS .....</b>                              | <b>XII</b> |
| <b>VITA .....</b>                                         | <b>XV</b>  |
| <b>CHAPTER 1: INTRODUCTION.....</b>                       | <b>1</b>   |
| Statement of the Problem .....                            | 1          |
| Study Purpose and Research Questions .....                | 2          |
| A Framework for Understanding “Data” .....                | 3          |
| Understanding Data “Use” .....                            | 6          |
| Study Significance and Implications .....                 | 7          |
| Manuscript Organization .....                             | 9          |
| Preview of Key Findings.....                              | 10         |
| <b>CHAPTER 2: REVIEW OF RELEVANT LITERATURE.....</b>      | <b>20</b>  |
| Introduction .....                                        | 20         |
| Current Literature .....                                  | 21         |
| <i>Processes of Data Use</i> .....                        | 21         |
| <i>Organizational and Political Context</i> .....         | 25         |
| <i>Interventions to Promote Data Use</i> .....            | 29         |
| <i>Potential Outcomes</i> .....                           | 32         |
| The Current Study .....                                   | 34         |
| <b>CHAPTER 3: RESEARCH METHODS.....</b>                   | <b>37</b>  |
| Introduction .....                                        | 37         |
| Study Procedures .....                                    | 37         |
| <i>Study Setting – Pilot Schools</i> .....                | 38         |
| <i>Comparative Case Study</i> .....                       | 39         |
| <i>Pilot High School Teacher Survey</i> .....             | 48         |
| <b>CHAPTER 4: SCHOOL DATA SYSTEMS AND STRUCTURES.....</b> | <b>50</b>  |
| Introduction .....                                        | 50         |
| Case #1: The Academy.....                                 | 52         |
| Case #2: Belleworth School of Arts and Technology .....   | 56         |

|                                                                                                      |            |
|------------------------------------------------------------------------------------------------------|------------|
| Case #3: Woodson College Preparatory School.....                                                     | 60         |
| Cross-Case Insights .....                                                                            | 67         |
| <b>CHAPTER 5: CULTURES OF DECISION-MAKING.....</b>                                                   | <b>69</b>  |
| Introduction .....                                                                                   | 69         |
| The Academy: Real-Time Decisions and Aspirations of Data Use .....                                   | 70         |
| <i>Decision-Making: Form vs. Function</i> .....                                                      | 71         |
| <i>Disparate Data Use Activities</i> .....                                                           | 76         |
| Bellewirth School of Arts and Technology: Power, Authority, and Then, Data.....                      | 78         |
| <i>Looking for Leadership in Data Use</i> .....                                                      | 78         |
| <i>Learning How to Leverage Data</i> .....                                                           | 80         |
| <i>Devolution, Dissolution, and Discord</i> .....                                                    | 83         |
| Woodson College Preparatory: Causal Relationships Rooted in Personal Relationships....               | 87         |
| <i>Internal and External Perceptions of Data Use</i> .....                                           | 87         |
| <i>Building Rapport and Gaining Perspective</i> .....                                                | 90         |
| <i>Securing Allies and Finding Pressure Points</i> .....                                             | 92         |
| Cross-Case Insights .....                                                                            | 93         |
| <b>CHAPTER 6: CULTURES OF “CREDIBLE DATA” .....</b>                                                  | <b>97</b>  |
| Introduction .....                                                                                   | 97         |
| The Academy: Data That Defines School Culture .....                                                  | 101        |
| <i>Measuring School Vision and Mission</i> .....                                                     | 101        |
| <i>Student Review Panels and the Complexity of Evaluating Academic and Behavioral Progress</i> ..... | 104        |
| <i>Innovations in Measuring Teacher Performance</i> .....                                            | 111        |
| Bellewirth School of Arts and Technology: Acknowledging Current Teacher Data Practices.....          | 123        |
| <i>Community-Based Intelligence</i> .....                                                            | 124        |
| <i>Building Student Rapport as a Means of Identifying Learning Strengths and Needs</i> ....          | 128        |
| <i>Student Data as Contributive to, Rather Than Predictive of Achievement</i> .....                  | 132        |
| <i>Teacher Interpretation of Student Statistics</i> .....                                            | 133        |
| Woodson College Preparatory School: Credible Data Is Meaningful Data .....                           | 141        |
| <i>Observational Data – Up Close and Personal</i> .....                                              | 142        |
| <i>Affective Data – More Than a Feeling</i> .....                                                    | 146        |
| <i>Enhancing Intuition</i> .....                                                                     | 148        |
| <i>Grades Ain’t Nothin’ But a Number</i> .....                                                       | 150        |
| Cross-Case Insights .....                                                                            | 155        |
| <b>CHAPTER 7: CULTURES OF DATA USE .....</b>                                                         | <b>158</b> |
| Introduction .....                                                                                   | 158        |

|                                                                                                               |                |
|---------------------------------------------------------------------------------------------------------------|----------------|
| <b>PART I: DATA USE IN STRATEGIC AND INSTRUCTIONAL PLANNING.....</b>                                          | <b>159</b>     |
| Bellewirth School of Arts and Technology: Using Data to Guide Program Development and Strategic Planning..... | 160            |
| <i>Using Data to Inform Student Supports and Interventions</i> .....                                          | 161            |
| <i>Forging Personal Connections With Data – A Prerequisite of Data Use</i> .....                              | 165            |
| Woodson College Preparatory School: Using Data to Guide Classroom Instruction .....                           | 170            |
| <i>The Science of Improvement</i> .....                                                                       | 170            |
| <i>Facilitating Constructive Conversations About Instruction Around Data</i> .....                            | 172            |
| <i>The Utility of PDSA Questioned as an Endless Cycle of Data Collection</i> .....                            | 177            |
| <i>When a Focus on Data Use Trumps Good Instruction</i> .....                                                 | 181            |
| <i>Woodson’s Identity Crisis</i> .....                                                                        | 182            |
| Cross-Case Insights .....                                                                                     | 188            |
| <br><b>PART II: DATA USE IN ASSESSMENT AND INSTRUCTION AT WOODSON COLLEGE PREPARATORY SCHOOL.....</b>         | <br><b>190</b> |
| The Common Assessments .....                                                                                  | 191            |
| The English Department: Assessments and “The Hidden Curriculum” .....                                         | 192            |
| The Science Department: Aligning Standards, Measures, and Instruction .....                                   | 196            |
| The Social Studies Department: Misalignment and Disenchantment .....                                          | 202            |
| Cross-Participant Insights.....                                                                               | 210            |
| <br><b>PART III: DATA USE IN SCHOOL PERFORMANCE MONITORING – IMPOSITIONS ON TEACHER AUTONOMY .....</b>        | <br><b>216</b> |
| Teacher Autonomy: Freedom, Power, and Duty.....                                                               | 217            |
| Something Borrowed, Something New: Teacher Buy-In, Ownership, and Ego.....                                    | 220            |
| <i>The Academy: Adaptation vs. Fidelity</i> .....                                                             | 221            |
| <i>Woodson College Preparatory School: The Expense of “Ownership”</i> .....                                   | 223            |
| Bellewirth School of Arts and Technology: Enforcing Standards of Success .....                                | 228            |
| Public Accountability .....                                                                                   | 233            |
| A Parallel Universe: District-Level Oversight and School-Level Discretion .....                               | 240            |
| Cross-Case Insights .....                                                                                     | 245            |
| <br><b>CHAPTER 8: THE STRENGTH OF THE ANECDOTAL: PROFESSIONAL JUDGMENT AS “SECOND TIER” EVIDENCE .....</b>    | <br><b>254</b> |
| Introduction .....                                                                                            | 254            |
| Why Art? .....                                                                                                | 255            |
| <i>The Classroom Play-By-Play</i> .....                                                                       | 255            |
| <i>Impressions as Imprints</i> .....                                                                          | 257            |
| <i>Assessing Assessments</i> .....                                                                            | 259            |
| <i>Outside Opinion</i> .....                                                                                  | 260            |
| Why Science? .....                                                                                            | 264            |
| Cross-Case Insights .....                                                                                     | 269            |

|                                                                                                                   |                |
|-------------------------------------------------------------------------------------------------------------------|----------------|
| <b>CHAPTER 9: DATA FOR ORGANIZATIONAL LEARNING VS. DATA FOR ACCOUNTABILITY .....</b>                              | <b>271</b>     |
| Perceptions of Data Misuse .....                                                                                  | 273            |
| Understanding Data in Context .....                                                                               | 280            |
| Practical Concerns, Conceptual Limitations .....                                                                  | 283            |
| Tainted Love.....                                                                                                 | 286            |
| Cross Case Insights .....                                                                                         | 292            |
| <br><b>CHAPTER 10: DISCUSSION .....</b>                                                                           | <br><b>295</b> |
| Introduction .....                                                                                                | 295            |
| Understanding and Supporting Data Use as a Part of School Culture .....                                           | 296            |
| Re-thinking Data Use for Decision-Making in Schools: A Revised Conceptual Framework .....                         | 299            |
| <i>Data Needs and Purposes</i> .....                                                                              | 301            |
| <i>Stakeholder Perspectives</i> .....                                                                             | 302            |
| <i>Decision-Makers and Decision-Making Processes</i> .....                                                        | 303            |
| <i>Data Systems and Structures</i> .....                                                                          | 304            |
| <i>The Identification of Credible Data</i> .....                                                                  | 305            |
| <i>Organizational and Individual Processes of Data Use</i> .....                                                  | 306            |
| <i>Practical Applications of a New Theoretical Approach</i> .....                                                 | 309            |
| Lessons Learned .....                                                                                             | 310            |
| <i>The Myth of Data Transparency</i> .....                                                                        | 310            |
| <i>Data Used in Decision-Making Are Part of the Process, Not the Outcome</i> .....                                | 312            |
| <i>How Data Are Not Used</i> .....                                                                                | 314            |
| <i>Treating Classrooms as Laboratories</i> .....                                                                  | 316            |
| <i>Professional Development</i> .....                                                                             | 319            |
| Study Limitations .....                                                                                           | 321            |
| Conclusion.....                                                                                                   | 322            |
| <br><b>APPENDICES .....</b>                                                                                       | <br><b>329</b> |
| Appendix A: Case Study Coding Framework .....                                                                     | 329            |
| Appendix B: Guiding Questions for School Leaders in Supporting the Effective Use of Data in Decision-Making ..... | 332            |
| Appendix C: Teacher Interview Protocol (Semi-Structured).....                                                     | 335            |
| Appendix D: Principal Interview Protocol (Semi-Structured) .....                                                  | 339            |
| Appendix E: District Personnel Interview Protocol (Semi-Structured) .....                                         | 344            |
| <br><b>REFERENCES.....</b>                                                                                        | <br><b>346</b> |

## List of Tables

|                                                                                   |           |
|-----------------------------------------------------------------------------------|-----------|
| <b>TABLE 1: CASE STUDY SCHOOL PARTICIPANT CHARACTERISTICS .....</b>               | <b>42</b> |
| <b>TABLE 2: CASE STUDY TEACHER AND PRINCIPAL PARTICIPANT CHARACTERISTICS.....</b> | <b>43</b> |
| <b>TABLE 3: INTERVIEW AND OBSERVATION DETAILS .....</b>                           | <b>45</b> |
| <b>TABLE 4: DISTRICT INTERVIEW DETAILS .....</b>                                  | <b>46</b> |
| <b>TABLE 5: DISTRICT OBSERVATION DETAILS .....</b>                                | <b>46</b> |
| <b>TABLE 6: DATA TYPES AND SOURCES REFERENCED BY STUDY PARTICIPANTS .....</b>     | <b>98</b> |

## **List of Figures**

|                                                                                                  |            |
|--------------------------------------------------------------------------------------------------|------------|
| <b>FIGURE 1: THE PROCESS OF TRANSFORMING DATA INTO KNOWLEDGE<br/>(ADAPTATION OF ACKOFF).....</b> | <b>5</b>   |
| <b>FIGURE 2: FRAMEWORK FOR DATA USE IN SCHOOLS (COBURN &amp; TURNER, 2011).....</b>              | <b>21</b>  |
| <b>FIGURE 3: CHALLENGES ASSOCIATED WITH MULTI-PURPOSE DATA<br/>(A TEACHER PERSPECTIVE).....</b>  | <b>291</b> |
| <b>FIGURE 4: FRAMEWORK FOR DATA USE IN SCHOOL DECISION-MAKING .....</b>                          | <b>300</b> |

## Acknowledgments

I have no less than tremendous appreciation for the mentorship, guidance, and unrivaled support of my advisor, Tina Christie, who has helped me to shape my time in graduate school into an incredible experience of scholarship in evaluation. Alongside her encouragement to challenge myself in the acquisition of new research skills, Tina has been a rock in my advancement to new parenthood. She has always backed my career aspirations and academic ambitions while fully understanding the practical demands of life. She has taught me that we *can* accomplish it all if we give one another the right support. For all of these things, I am endlessly grateful.

Great recognition is owed to the Los Angeles Unified School District (and in particular Kathy Hayes), for its support of this research and its willingness to engage in constructively critical dialogue. To the tireless, dedicated members of the former Intensive Support and Intervention Center (ISIC), who took me under their wings in coming to understand their work and their schools, thank you for your trust, your endorsement, and for your time in conducting this research. I am also indebted to each of the principals, teachers, parents, and school faculty members participating in this study, who have gifted me with their candid honesty, incredible insight, and patient detail of their experiences and practice. Your perspectives are the real substance of this research, and I cannot thank you enough for so graciously inviting me into your professional spaces.

This piece of work could not have been accomplished without the investments of my own teachers. To Marv Alkin, thank you for integrating my thinking into evaluation use theory – the stuff that guides my work daily – and for all that you have done to help me become a critical scholar. To Katie Anderson-Levitt, I have endless gratitude for the generosity with which you

have offered your listening ear, your expert consultation, and your gentle guidance as I have come to understand qualitative research over these past five years. To Bruce Fuller, your willingness to take a chance on me, and to continuously press me into considering what my work means to schools in the language of schools is not without great acknowledgment. To John Rogers, thank you for helping me to navigate the complexity that is school politics, and for helping me to root my work in the context and needs of our community. Great thanks and appreciation are very much owed to Mike Rose, whose friendship continues to shape my thinking and writing long after class time is over, as well as to Todd Franke, for always finding a spare minute to help me understand what it is I am doing, why I am doing it the way that I am, and for your compassionate approach to those of us still learning. To Mike Seltzer, Reenie Webb, Felipe Martinez, Li Cai, Terri McCarty, and Mark Hansen, thank you for being not only excellent faculty but also wonderful people to work with and learn from.

My graduate school experience would have been dimly lit without the laughter, brilliance, and direction provided by my colleagues. To the evaluation students before me (Lisa, Tim, Debbie, Ale, Celina, and Nicky), to my cohorts (Megan, Jane, Danny, Liz, Kevin, and Jason), and to those current evaluation students (Sebastian, Minh, Adrienne, Alana, Kristen, and Talia), thank you for cheering me across the finish line, and better yet, for being your help and support all the way through – my sanity would have been lost long ago without you all. A special thank you to Patty for your constant reassurance, for reminding me when to slow down, when to laugh louder, and to take it all in stride.

My family members often bear the brunt of my stress, but this has never diminished their encouragement of my work or their pride in my accomplishments. Thank you Mom, Dad, Lorraine, and Josh for being frequent visitors to Los Angeles, for watching the kids when I

worked, and for always having confidence in my ability to get this done. To the village that is Pacific Street (Kristen, Zak, Bodhi, Erin, Dan, Nova, Nina, Ross, Maya and Jade), thank you for being my extended family and great friends – my indisputable lifeline.

I must also give thanks to the ongoing support I have received from my EDC, Inc. family, including access to data sets for research projects. Steve Anzalone, you remain an incredible mentor and someone I can always depend upon for a pertinent perspective on life. Thank you for your unending advocacy on this journey to obtain my “black belt,” even though it meant leaving your team to do so. My work with you and our time in Indonesia was more foundational than I could ever have realized.

Last, but never least, none of this would have been possible without the willingness of my husband, Roy, to move across the country and begin this incredible adventure, to be a willing sounding board for every piece of this dissertation, and to bring into the world not one, but two little boys over the course of this study. Your support has manifested itself in innumerable ways, but few were as important as you repeating, “It’s going to be all right,” as many times as it took until I believed you. Thank you, for believing in me.

# VITA

## EDUCATION

- 2005                      M.Ed. International Educational Policy  
Harvard Graduate School of Education, Cambridge, MA
- 2002                      B.A. Political Science/Philosophy  
Boston University, Boston, MA

## PROFESSIONAL EXPERIENCE

- 2015 – 2016              Graduate Student Researcher  
NIH Diversity Program Consortium Coordination and Evaluation Center  
University of California, Los Angeles
- 2014 – 2015              Research Specialist  
School-to-School International  
Pacifica, CA
- June – Sep 2013        Graduate Student Fellow  
The Broad Center  
Education Pioneers, Los Angeles
- 2012 – 2015              Graduate Student Researcher  
Center for Healthier Children, Families & Communities  
University of California, Los Angeles
- 2012 – 2013              Teaching Assistant  
Graduate School of Education and Information Studies  
University of California, Los Angeles
- 2006 – 2011              International Technical Associate  
Education Development Center  
Washington, DC

## SELECTED PUBLICATIONS

- 2012                      Ho, J. and Thukral, H. Interactive Radio Instruction as a Distance  
Education Approach in Developing Countries. In Lya Visser, Yusra

Visser, and Rya Amirault (Eds.), *Trends and Issues in Distance Education: International Perspectives, Second Edition* (113-124), Charlotte: Information Age Publishing.

- 2009                      Ho, J. and Thukral, H. *Tuned In to Student Success – Assessing the Impact of Interactive Radio Instruction for the Hardest-to-Reach*. Washington DC: Education Development Center, Inc. Published also in *Journal of Education for International Development (JEID)*, Volume 4, Issue 2, *ICT and Education*, December 2009.

## **SELECTED PRESENTATIONS**

- 2015                      Ho, J. *School Autonomy and Accountability: A study of how pilot schools use data to inform decision-making*. Paper was presented at the American Evaluation Association Conference. November 13, 2015.
- 2013                      Ho, J. *Measuring Changes in Teacher Practice: The use of multi-level modeling in international education evaluation*. Paper was presented at the American Evaluation Association Conference. October 17, 2013.
- 2013                      Ho, J. *Exploring The Adoption of Active Learning Techniques in Indonesian Classrooms: A longitudinal analysis of teacher practice using hierarchical linear modeling*. Paper was presented at the Comparative International Education Conference. March 13, 2013.
- 2012                      Ho, J. *The Mobile Gourmet: How food trucks swayed the popular palate and stirred the culture of gourmet cuisine*. Paper was presented at the American Anthropological Association Conference. November 16, 2012.

## **SELECTED HONORS AND AWARDS**

- 2015-2016                      Dissertation Year Fellowship, UCLA Graduate Division
- 2014-2015                      Graduate Research Mentorship Fellowship, UCLA Graduate Division
- 2012                              Graduate Summer Research Mentorship, UCLA Graduate Division
- 2011-2014                      William and Louise Lucio Fellowship, UCLA GSE&IS

# **CHAPTER 1**

## **INTRODUCTION**

### **Statement of the Problem**

The reliance of American education on test-based accountability policies to improve student achievement has been in practice since the 1970s. Despite these efforts, discrepancies in student achievement have been persistent and research on remedial education innovation has been characterized as lacking. Critics have argued that improvements in education have been compromised by practitioners' propensity to base change not on the progress of scientific inquiry and research-based evidence, but rather on the "pendulum swings of taste characteristic of art or fashion" (Slavin, 2002).

In response to the perception that educators ground their decisions in fallible instinct, intuition, and fad, U.S. schools have experienced a resurgence of accountability policies at both the federal and state levels. The No Child Left Behind Act (2002), followed most recently by the Every Student Succeeds Act (2015), the America Reinvestment and Recovery Act (2009), and the Statewide Longitudinal Data System and Grant Program (2005) are prime examples of a conversation re-focused on the use of data for purposes of school accountability and improvement. Leading funders in educational reform, such as the Stupski Foundation and the Gates Foundation, have also taken a prominent stance on the issue, the latter having pledged \$12 million to support investment and implementation in school data systems (Coburn & Turner, 2012). Under these policies and initiatives, there has been a large push for schools to engage in decision-making based upon empirical data; that is, schools are responsible for collecting data through observation and experimentation, and are also expected to incorporate these data into decisions made around teaching and learning. School districts have thus invested in data systems

in order to create enhanced access to data, as well as training teachers, principals, and district leaders to focus on the integration of data into their daily practice (Datnow et al., 2007; Kerr et al., 2006; Marsh et al., 2006).

Subsequently, “data-based decision-making” or “evidence-based decision-making” in education are now popularized phrases to describe the systematic collection and analysis of various types of data, including *input*, *process*, *outcome*, and *satisfaction data*, to help improve the success of students and schools (Marsh, Pane, & Hamilton, 2006). The theory of action underlying data use activities implies that education practitioners who ground their decisions in evidence will better ensure methodical improvements to teaching and learning. However, very little research has been conducted to test this hypothesis. In addition to the research community’s vague understanding of how schools – and the individuals comprising schools – interpret and implement data-based decision-making policies, it is difficult to determine whether data use practices are actually associated with improved instruction. As a result, school districts, as well as state and federal policy makers, have little understanding of how schools are actually using data, how differences in data use may affect school performance, and/or what kinds of measures could be used to indicate the effective use of data in schools.

### **Study Purpose and Research Questions**

This study focuses on the contextual factors influencing how data is defined and used to make specific decisions regarding policy and practice in a subsection of schools in the Los Angeles Unified School District (LAUSD). In so doing, it seeks to understand how data is construed and interpreted by local school stakeholders. It attempts to delineate the ways in which school stakeholders apply data in their naturally varying contexts and to explore how data are identified, valued, and used to influence decisions relative to other considerations. Through the

exploration of the ways in which data are applied – or are not applied – in the day-to-day functioning of schools, this research hopes to gain a better idea of what data use looks like from the perspective of schools and their various stakeholders.

The specific research question guiding this study asks, *how do teachers, principals, and district personnel use data in their professional contexts?* To address this overarching inquiry, several specific questions were pursued:

1. What do school practitioners identify as data, and particularly as credible data?
2. How do teachers and principals use data to inform decisions related to school improvement and strategic planning?
3. How do teachers use data to inform instruction?
4. How do teachers, principals and district personnel use data to monitor school performance?
5. How do organizational and cultural characteristics of schools affect the way teachers and principals use data (for any of those purposes)?

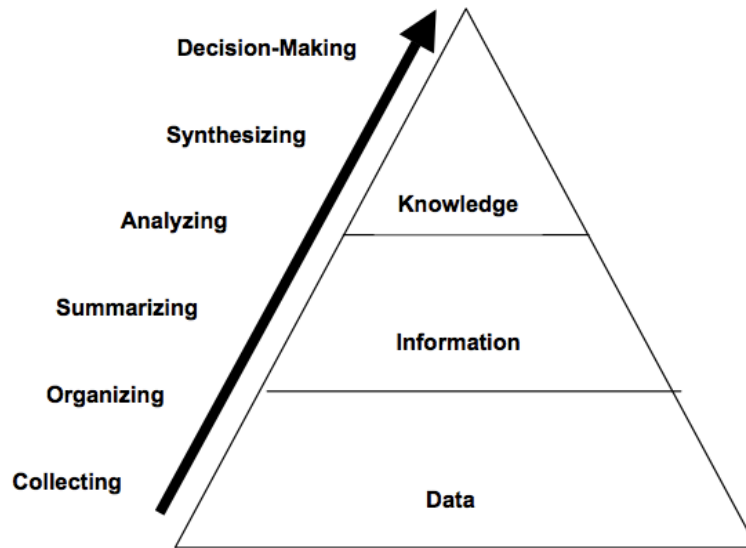
In addressing these questions, this study intends to develop a more concrete understanding of how school decision-makers (i.e., teachers, principals, and district personnel) make meaning of directives to “use data for decision-making” and how the use of school-based data takes place in practical application.

### **A Framework for Understanding “Data”**

The view of data undertaken in this research is broad in order to allow for participant interpretation. It includes not only the kind of data focused on previously-validated measures of student and school performance, such as student assessment results or graduation/attendance rates, but also takes into consideration what Coburn and Turner (2012) describe as “how people

use measures of social and organizational conditions and information that they gather through their experience” (p. 100). This study recognizes that data are not objective guides in making decisions but instead rely on practitioners’ abilities to identify and interpret their meaning. This study considers research that suggests good, applied practice is predominantly dependent on accumulated experience combined with local ideas, attitudes, and discussion (Wood, Ferlie, & Fitzgerald, 1998). It recognizes that teachers, administrators, and policy makers call into practice various sources of information drawn from experience and observation, not just social science research and student achievement data (Kennedy, 1982; Little, 2007). This take on data may thus include results of research and evaluation – distinct endeavors that each entail its own theoretical approach to “use” (Alkin, 2004; Nutley, 2007) – but is not restricted to the output of these activities.

In its raw form, “data” is treated separately in this study from “information” and “knowledge.” Ackoff’s (1989) well-known work in organizational and management theory proposes a “structure of knowledge” wherein *data*, *information*, *knowledge*, and *wisdom* are hierarchically arranged as ascending levels. In this framework, each of the categories includes the one below it (such that, for example, there can be no wisdom without understanding, and no understanding without knowledge). Adaptations of Ackoff’s (1989) framework in educational research, such as that proposed by Light et al. (2004), shown in Figure 1 below, reference the first three categories of this hierarchy.



*Figure 1: The Process Of Transforming Data Into Knowledge  
(Adaptation of Ackoff)*

From this standpoint, “data” do not have meaning in and of themselves and can exist in any form, usable or not. Whether “data” become “information” depends upon the understanding of the individual interpreting the data: “information” is described as data that is given meaning when connected to a context; it is data that are used to comprehend and organize our environment, and draws relationships between data and context (Ackoff, 1989). In this framework, information alone does not carry any implications for future action. “Knowledge” is the collection of information regarded as useful and is eventually used to guide action. This hierarchy of knowledge is described as necessarily sequential, such that in order for teachers or administrators to make knowledgeable decisions about teaching and learning, they must first be able to identify a data source and collect and organize that data. Data must then be analyzed and summarized; data become information when their meaning is interpreted alongside other sources of various data. Finally, to turn information into knowledge, stakeholders must synthesize all of

the available information and place a value judgment on that information through prioritization. This process entails the determination of the relative importance of information and the consideration of possible actionable solutions.

Proponents of data use in school-based policy development and decision-making have used the phrase “evidence-based decision-making” interchangeably with “data-based decision-making.” It should be noted that “evidence” and “data” are treated as distinct terms in research literature, where “evidence” is considered “a value-based label attached to particular types of knowledge” (Nutley, 2007). However, these two phrases are regarded as the same in their intent and in their description of school decisions founded on empirical information.

### **Understanding Data “Use”**

Research and theory point to several different types of data “use.” The language of education reform initiatives focuses primarily on the use of data for the direct purpose of decision-making. Cousins and Leithwood (1986) define this type of use as a “discrete activity related to decisions about program funding, the nature or operation of a program, or regarding program management.” However, they also identify several other types of data use relevant to schools including use as *education* (i.e., the enlightenment of decision makers by influencing their perceptions of current and ideal program structures), the simple *processing of evaluation results* (i.e., when findings have been given some thought or consideration, including basic understanding of evaluation data), and the *potential* for use (i.e., users’ satisfaction with evaluation recommendations and estimated influence on future decisions). It is recognized that use may not only be *instrumental*, such that observable action can be definitively linked to data, but also *persuasive*, wherein individuals use data to support their own positions and beliefs for personal or political gain. Use may also be *conceptual*, wherein data may influence individuals’

thinking about a program or issue (King, 1988; Leviton & Hughes, 1981). The very process of using data is described by Patton (2008) as influential in helping individuals and organizations to “think evaluatively,” and for the latter to become “learning organizations.” The “non-use” or even the “misuse” of data are also important elements in understanding when data are justifiably or unjustifiably, intentionally or unintentionally, neglected, suppressed, or abused in its consideration (Alkin & Coyle, 1988; Patton, 1988). Examples include the commissioning of evaluation for purely symbolic reasons, the conscious subversion of evaluation by program practitioners, and the purposeful non-use of high-quality information (King, 1988). Data can also be used as “instruments of persuasion” to mobilize support for a position people already hold about the changes needed in a program (Weiss, 1998). Given all of these distinctions, the question of how exactly data are used in school contexts is as consequential as how data are identified and defined in practice. How “use” is interpreted thus remains a prevalent question in understanding how schools respond to the promotion of “data-driven decision-making” as a best practice.

### **Study Significance and Implications**

Research to date has thus far indicated a number of components critical to functioning systems of data based decision-making within schools. These include resources (such as time, technical expertise, and an infrastructure through which data are accessible), school cultures supportive of inquiry and trust among colleagues, and school-based policies guided by visionary leaders with a commitment to data use. While the identification of these elements is an important contribution to our understanding of what is needed to support the use of data in decision-making, they are often regarded as inputs that can be introduced or enhanced to improve school-based data use.

Missing from the conversation is a more complex understanding of the role of cultural development in shaping data use processes and outcomes. “Culture” is referenced within this research as a shared social meaning constructed from the common experiences of individuals. From this perspective, effective data use is understood as one objective among many within a school. As schools develop, revise, and refine processes of data use for a variety of purposes, contextual factors are perceived to influence what schools - as a collective unit - identify as data, what they prioritize as valuable data, and in what ways they make use of data (if at all). What schools glean from their data, and their experiences participating in data use processes, may, in turn, affect stakeholders’ approaches to decision-making: who takes part in decision-making and what decisions are eventually made. This notion of *cultural influence* extends well beyond one of inquiry or collegiality encouraging of honest discussion, analysis, and interpretation in a cooperative response to data. Rather, it takes into consideration the broader aspects of school-based decision-making and the ways in which culturally-derived definitions of credible data, leadership, decision-making processes, accountability, organizational learning, and evaluation – and even whether data are relevant in teachers’ thinking in institutional contexts – shape stakeholder attitudes toward data use in classrooms and schools.

By understanding what cultural factors underlie data use in schools, and the ways in which they develop and unfold, we gain a better perspective of not only what schools need to support effective data use but, more importantly, what this looks like in implementation. This study approach intentionally acknowledges the work that is currently being accomplished by schools in their use of data, as well as the complexities they confront in doing so. The voices of schools and school stakeholders are critical in the conversation about data use in schools because, at the end of the day, data are targeted at the improvement of teaching and learning.

Data are used as indicators of effective instruction and are ultimately expected to guide teachers and administrators in making instructional changes supportive of improved student achievement. By teasing apart potential discontinuities between how data are used in practice at the school- and classroom-levels, as well as expectations of data use implied at the policy-level, this study sheds light on how organizational and instructional change – rooted in context – both drives and is driven by concepts and processes of data use.

### **Manuscript Organization**

In exploration of the cultural and contextual influences on data use in school-based decision-making, Chapter 2 presents a review of literature relevant to our current understanding of data-based decision-making in schools, and Chapter 3 details the methods of research employed within this study. Chapter 4 presents the three case study sites in a discussion of the systems and structures underlying their differential use of data, and Chapter 5 provides the context for how decisions are made within each school site. Chapter 6 details the various types of data stakeholders within each case deem credible in their practice. Chapter 7 consists of three parts in its discussion of how data are used within schools: Part I reviews how, in two cases, data are used to inform instructional and strategic planning; Part II looks intensively at one school in its use of student assessment data to inform instruction; and, Part III discusses the use of data to inform school performance monitoring and how this interacts with notions of teacher autonomy. Chapters 8 and 9 present themes resulting from cross-case analyses; Chapter 8 pays particular attention to the value of anecdotal data in assessing student and school performance, and Chapter 9 looks closely at issues arising from the use of data for both purposes of accountability and organizational learning. While Chapter 10 provides a more detailed discussion of study results,

key findings from each chapter are provided below as a precursor to the in-depth analyses provided within each chapter.

### **Preview of Key Findings**

1. *Systems and structures of data access, review, interpretation, and use were important, not imperative.*

Systems and policies of data review, as well as organizational structures promoting data routines, are presented in this study as an underlying feature of data use within each school case. The development of each pilot high school – from concept to implementation – as well the constitution of its faculty, governance structure, and its maturation of mission and vision, are all seen to contribute toward a school’s active use of data in decision-making. Chapter 4 explores whether each school has taken stock of, and amassed, various data sources to which it has access, as well as whether schools have introduced procedures of data use, including determining who will be included in data conversations, regularly scheduling reviews of data, and designating time for those reviews. The chapter begins to outline each school’s intentions in using data for decision-making and the level of practice in translating conversations around data into conversations around actionable next steps. Variation of these many factors within each school case suggests that the direct comparison of data use “proficiency” across schools is not as appropriate as understanding data use as context-dependent. Indeed, it was found that schools can and do use data, even when formal data routines and infrastructure to support data compilation and analysis are not yet established.

2. *Transparent processes of decision-making and the authentic engagement of school stakeholders in decision-making were prerequisites to data use.*

The types of decisions involving data range widely among schools. Examples from this study include the use of data by school leadership to inform the development of student support interventions, as well as the use of data by teachers in moments of instruction. In the discussion of whether and how schools are using data for decision-making, it is important to recognize that schools are not single entities, but rather units comprised of multiple stakeholder groups.

Students, parents, teachers, principals, and District administrators are all seen to be agents of data use at the school-level. Within those groups, individuals bring to bear their own perspectives, priorities, and values to the decisions made on their campus. However, decision-making processes are not necessarily all-inclusive. Rather, as Chapter 5 illustrates, the degree to which decision-making processes are established and entrenched, and the ways in which various stakeholders are actively incorporated into those processes, were found to largely determine the degree to which data were referred. In the three cases observed, the determination of what kind of data should inform decision-making was not as much of a priority as what decision-making processes would dictate data use. This study has shown that who determines what should be done with school data substantially influences whether and what data are actually referenced in making decisions. This is not merely a designation of responsibility or even just an issue of authority; rather, stakeholders' perceived senses of value as decision-makers and their control over decision-making processes were observed to impact genuine engagement. Systems supportive of collaboration, open dialogue, transparent negotiation, process-oriented decision-making, and a common vision toward teaching and learning are regarded as prerequisite to the consideration and subsequent incorporation of data into decision-making.

3. *Data credibility was context- rather than criteria-dependent.*

Acknowledgment of the many individuals comprising schools also lends itself to the exposition of the assorted perspectives contributing to definitions of “credible data.” Chapter 6 examines what data participants consider reliable, relevant, and accurate in responding to questions about student learning and teacher instruction. In many circumstances, the types of credible data prioritized by teacher participants fell outside criteria for systematically-collected school-based data commonly referenced by proponents of data-based decision-making. That is, cited sources of “credible data” were not always drawn from the category of routinely-collected, systematically-reviewed, and collaboratively-assessed and interpreted data (such as student outcome data). Instead, teachers were found to frequently rely on observational data related to student academic achievement and behavior, student background and contextual data, and anecdotal data indicative of student improvement as pieces most relevant to their own instructional moves. This is not to say that data sources, such as student outcome data, were not of value – teacher participants frequently endorsed these data for purposes of accountability and drew on these data for use in instruction when appropriate. However, teachers did feel that the types of data they personally found most useful were not always recognized as “credible” in external evaluations of student and school performance. Taken together, these findings suggest that data credibility is not objectively conferred as a veritable truth but, rather, that data gain and lose credibility in their applications to different purposes.

4. *Data, data collection, and expectations for data use needed to be aligned with instruction and instructional needs.*

Alongside the discussion of what data are considered credible is the articulation of how data are actually used in processes of school decision-making. Chapter 7: Part I begins to unpack

how data are folded into conversations around strategic and instructional planning, student assessment and instruction, and school performance monitoring. While multiple examples of data use are discussed throughout the study, examples from Belleworth School of Arts and Technology<sup>1</sup> are drawn upon to illustrate the ways in which the analysis of student performance data can contribute to the development of student support programming and the identification of students needing intervention or enrichment services. Data use within Woodson College Preparatory School is explored through teachers' implementation of the Plan-Do-Study-Act initiative (PDSA) designed to guide teachers through their own collection, interpretation, and application of data in refining classroom pedagogy. For both schools, it is clear that teachers' sense of connection and responsibility to data are essential to making use of that data. This includes the ability to understand that "numbers," presented by student performance data for example, reflect actual students affected by teachers' classroom practices. Indeed, structured discussions about content and curriculum using data collected by teachers are observed to foster productive conversations about instructional strategy and pedagogical approach. However, it is also observed that teacher ownership of data can be interpreted by teachers as a burden. Teacher-collected data can feel overwhelming, exhausting, and pointless when data collection procedures are not well-aligned with the flow of everyday classroom procedures, when teachers are unclear about what types of data constitute rigorous examinations of teacher practice and student learning, and when facilitators of data use processes do not acknowledge the intensive resources required to effectively interpret data into instructional change. As a result, while teachers may conceptually endorse the use of data to make decisions around school and instructional strategy, doing so does not necessarily translate into the actual application of data for these purposes.

---

<sup>1</sup> Pseudonyms for participant schools were used to protect participant identity.

5. *Student assessment data were more likely to be used when teachers were actively engaged in test design, implementation, and scoring, and were given the opportunity to reiterate cycles of test development.*

The experience of designing and developing student assessments at Woodson College Prep is explored as an example of what *complete teacher ownership* of data collection and use processes looks like in implementation in Chapter 7: Part II. The differential immersion of Woodson's English, science, and social studies departments into processes of test construction and scoring presents three diverse pictures of student assessment data use. All of the departments found that student test development takes time, not just in terms of item construction or the identification of an appropriate scoring rubric, but also in terms of repetitive cycles of implementation. Observation and analysis of how students interact with assessment content, whether and how students' skills and abilities are elicited by test items, and how scoring criteria are applied to student work serve as conduits for teacher conversation around prioritized student learning outcomes, indicators of academic progress, and plans to further support student achievement. They allow for both teacher reflection on what students know, as well as whether assessments and scoring criteria adequately capture student ability. The constant exchange between processes of test development and data interpretation ensures teachers' essential role as translator between assessment results and instructional change. *Teacher capacity building* in student assessment is necessarily experiential as teachers work through how assessments react to changes in student performance and vice versa. On the contrary, it was found that teachers' detachment from processes of test development, scoring, and analysis could result in a great deal of misunderstanding around how tests are best conducted and what value they hold for instruction. The use of assessment data to improve student learning, then, is dependent on

teachers' working knowledge of testing procedures and the direct correlation of test content to instructional content.

6. *Teacher "buy-in" into data use processes was distinct from teachers' sense of "proprietary ownership" of data use processes.*

In Chapter 7: Parts I and II, it was observed that data are more likely to be used in classrooms when teachers have a sense of ownership over the ways in which data are derived and interpreted. However, in Chapter 7: Part III, investigating schools' experiences using data for school performance monitoring found that teacher ownership over data use processes can sometimes be regarded as being "proprietary" rather than "involvement" or "endorsement." That is, some teachers seemed to need complete jurisdiction of all data use processes, at times endangering the rigor or methodological strength of data collection plans and procedures. A careful balance was also observed between teachers' perceived autonomy over the use of data in their school and the establishment of a culture of mutual accountability among school stakeholders. In some spaces, teachers were seen to push one another to higher levels of performance through constructive conversations around student outcome data. These teachers also actively participated in the collective development of learning standards to which they hold one another accountable. In other spaces, teachers were reluctant to share their data with colleagues, or acknowledge school-based data as a reflection of student and teacher performance. This is partially discussed as an element of control, wherein teachers felt the need to withhold data because of their concern in not accurately portraying student knowledge or the effects of their own instruction. It is important to recognize data limitations and that teachers and administrators cannot realistically control all factors influencing performance data. However, reticence to view student outcome data as a measure of school performance has also been

discussed as an issue of “ego.” Some participants suggested that teachers may need to relinquish their territorial grasp on data in order to learn from them, and that seemingly negative results should be approached with humility, understanding, and a determination to improve. Even though student assessment data are recognizably imperfect, it is suggested that these data still provide essential metrics of a school’s effectiveness in serving students.

7. *Anecdotal evidence was valued as credible data, particularly as informants to professional judgment.*

The bulk of this study focuses on what data constitute credible measures of student, teacher, and school performance, as well as the ways in which those data are or are not used in processes of decision-making. Chapter 8, however, dedicates some time to the consideration of a type of data that is considered extremely valuable by teachers and is used regularly in the course of their work, but which is regarded as auxiliary by those interested in the objective evaluation of schools. Specifically, the need for teachers to exercise professional judgment and make instructional decisions in-time with student learning necessarily incorporates *anecdotal evidence*. Anecdotal evidence is referenced by multiple teacher participants as a kind of data indicative of the experience of individual students as he/she undergoes processes of learning. These data are seen to feed teachers’ intuitive responses to the ways in which students grapple with curricular material and their progression as critical thinkers and learners. Importantly, they help teachers generate hypotheses about their instructional practice. Anecdotes shared among teachers inspire reflective questioning as to what implications students’ learning experiences in different scenarios have in their own classrooms. As teachers glean bits and pieces of anecdotal data through teaching and learning transactions, these data contribute to a larger body of evidence in the consideration of student and classroom-related instructional issues. Anecdotal evidence often

pertains to performance outcomes that are difficult to empirically measure and for that reason, are also viewed as part and parcel of assessing the success of schools in meeting the needs of their students and communities.

Anecdotal data are not necessarily considered an infallible foundation of understanding what goes on in the classroom. One teacher participant suggested that the more systematic collection and analysis of classroom data is a worthwhile endeavor and may indeed improve teachers' accuracy in determining how their students might more effectively engage in learning. Additional teacher participants suggested that anecdotal evidence should not be the sole source of data on which to wholly assess student progress. While the limitations of anecdotal data are recognized, there exists a need to dignify the necessary role they play in guiding teacher action and to protect teachers' discretionary use of judgment as education professionals.

8. *Data became less understood if they were used for multiple purposes.*

Chapter 9 more specifically addresses the issue of leveraging sources of school-based data for multiple purposes. Participants were more likely to express misgivings about data credibility and use when data were designated to serve multiple purposes or when the motivations guiding data use were unclear. Indeed, the use of data for unanticipated purposes can result in substantial ramifications. Teachers and principals alike have seen, for example, how seemingly benign data have been misused for political leverage or manipulated by schools for purposes of reputation and/or gain. Experiences like these contribute to stakeholders' wariness of data, and at times, resentment over the power data can wield in high-stakes decision-making. Data are considered to be particularly insidious when analyses ignore contextual factors. Teacher and principal participants consistently emphasized the importance of decision-makers'

understandings around how student and school performance data are composed, as well as the many factors contributing to their fluctuation and variation. Equally as critical is the recognition of data limitations. Single pieces of data are naturally confined depictions of educational outputs and outcomes – they do not capture the entire complexity of teaching and learning processes.

Those outside of the classroom are encouraged to recognize that data are a naturally delimited portrayal of instructional efforts. On the other hand, data are considered important indications of student progress, as well as teacher and school effectiveness, and teachers fulfill an essential role in translating performance “numbers” into instructional improvement. Mutual understanding among in- and out-of-classroom stakeholders, however, is particularly complicated when school-based data are used for both purposes of organizational learning and school accountability, a dichotomy frequently encountered by teachers and administrators. While data used to inform organizational learning are meant to contribute to a school’s continual improvement as defined by internal standards of success, the need to respond to external expectations of achievement orient data use toward compliance standards. One teacher provided an example of how data she considered extremely useful for her instruction could become stigmatized when it was also published as an indicator of school effectiveness. The pressure to evidence improvements in student performance, she argued, led her focus away from individual student progress (inherently varied in pace and substance) and toward a strategically-designed progression through the curriculum. This can lead to teacher frustration with students and demoralization when data goals are not met. Yet, however unintentional, data that are genuinely used to improve teaching and learning – within and between classrooms – irreversibly lose their integrity when co-opted as performance metrics. This is not to say that schools have no responsibility to produce accountability data; but in consideration of how to promote authentic

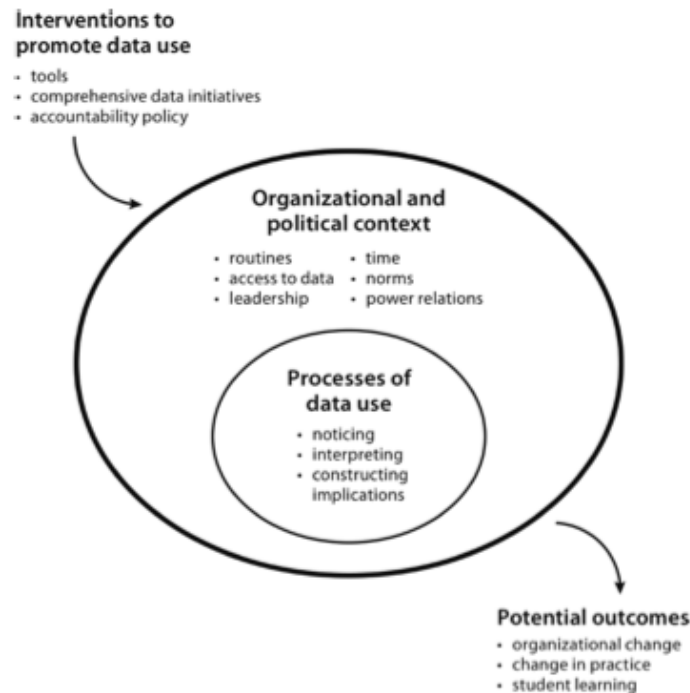
data use in schools, it is argued that researchers, evaluators, and policymakers must clearly communicate their intentions in using school-based data and honor these agreements with schools. Perhaps even more important, it is imperative to acknowledge the unintentional consequences involved in re-purposing data when it occurs.

## **CHAPTER 2**

### **REVIEW OF RELEVANT LITERATURE**

#### **Introduction**

While not prolific, research on the use of data in schools for the purposes of decision-making has grown in response to policy mandates and other funded initiatives that encourage data use. Though many influential frameworks for understanding data use in practical contexts exist, few of them specifically address data use in schools. One particular framework, however, is instrumental in organizing the corpus of theoretical work surrounding data use practices in schools and is presented by Coburn and Turner (2011). Attempts to apply “use” typologies to practice has shown that the use of data is, in fact, a dynamic process: different types of use interact and build on one another more often than behaving linearly (Nutley, 2007). The Coburn and Turner (2011) framework begins to acknowledge this fluidity, as well as the influence of social contexts and power relations on data use activities. Importantly, it treats the interpretation of data and its use as a complex undertaking intimately linked with social, political, and procedural pressures. The anticipated outcomes of these efforts are, ultimately, improved teaching, learning and organizational change. This comprehensive view (see Figure 2), combined with its thorough review of current literature on school-based data use, is what makes the Coburn and Turner (2011) framework an especially instrumental orientation for this study. However, its regard of schools as formal decision-making structures – and decision-making as a logical process undertaken by groups of rational decision-makers – presents a narrow view to schools’ use of information in light of organizational and decision-making theory.



*Figure 2: Framework For Data Use In Schools  
(Coburn & Turner, 2011)*

## Current Literature

### Processes of Data Use

The center of the Coburn and Turner (2011) framework depicts the “process of data use” which they define as what actually happens when individuals interact with assessments, test scores, and other forms of data in the course of their ongoing work (p. 176). In alignment with Ackoff’s “structure of knowledge” (1989), Coburn and Turner note that data use is an “interpretive process that involves noticing data in the first place, making meaning of it, and constructing implications for action” (p. 175). As an inherently interpretive process, data use processes are explained as subject to the characteristics of the individuals involved and the dynamics of their social interaction with others.

A good deal of research suggests that what data teachers eventually use depends on what data are considered “credible,” and that what teachers identify as *credible evidence* is often influenced by what matches their personal experience (Zeuli, 1994). As a result, what data users search for and see in the data largely depends on what findings support their own beliefs, assumptions, and experiences (Bickel & Cooley, 1985; David, 1981; Donaldson, Christie & Mark, 2014; Hannaway, 1989; Ingram, Seashore Louis, & Schroeder, 2004; Kennedy, 1982; King, 1988; Rickinson, 2005; Weiss, 1995; Young & Kim, 2010). In some cases, individuals may not even notice data that may contradict their beliefs. This is especially true where educators are forced to narrow the range of what they search for and pay attention to when schools are inundated with data (Honig, 2003). Similarly, it has often been found that the interpretation and application of data also relies upon a series of individual assumptions, conjectures and judgments rooted in one’s prior beliefs and experiences (Weiss, 1999; Court & Young, 2003; Kennedy, 1982; Little, 2007). Simons and colleagues (2003) warn that traditional notions of validity may not necessarily apply in education wherein teachers are swayed both socially and situationally, often judging the value of research for their practice based on other teachers’ assessments of research and its usefulness.

In addition to the credibility of data, there is also the consideration of what is relevant to school stakeholders. A separate framework put forth by (Gill, Coffee-Borden, & Hallgren, 2014) draws a distinction between valid and reliable data (necessary elements for the use of data diagnostically in education settings), and data that is relevant. In this framework, student, staff, and school or program data relevance is dependent on the different needs of classroom teachers, school administrators, central office administrators, and state education officials. How often data

need to be updated, and the level of detail in the data provided, are presented as key elements of relevance that vary between each stakeholder group.

Also not featured within the Coburn and Turner (2011) framework – and perhaps underlying all of these data use processes – is the importance of individual capacity in analyzing, interpreting, and manipulating data. The effective, efficient, and reflective use of data to drive decision-making is an activity described by Mandinach, et al. (2006) as one influenced by more than technological tools and general human capacity development. More specifically, the interpretation of data begins with one’s foundation in basic statistical concepts. The ability to move beyond interpretations of individual student performance to the description of aggregate student behavior, for example, requires an understanding of distribution, sampling, variation, and statistical difference. Being able to differentiate individual student performance from “averaged” results (such as in the identification of “high-risk” students), requires that educators have an understanding of *variation* and *distribution*. Examining differences between student groups also requires an understanding of what constitutes “significant” variation between groups, as well as how to interpret interactions and when to investigate normal variability. As such, the interpretation of data and educators’ roles in implying action rely heavily on individuals’ statistical fluency. Beyond statistical interpretation, several studies point to the lack of capacity of school personnel to formulate questions, select indicators, and develop solutions (Dembosky, Pane, Barney, Christina, & Education, 2006; Mason, 2002; Quartz, Kawasaki, Sotelo, & Merino, 2014).

In relation to individuals’ perceived self-capacity in working with data, there is also the issue of individuals’ interest in, and propensity toward data use. Feldman and Tung (2001) expand upon the importance of individual facility with data in their observations of teachers

engaging in data-based decision-making initiatives. When a “lead” teacher from their case study, who was considered most comfortable and skilled in data interpretation, left his school, they observed that his colleagues did not believe they could complete their data-based inquiry project without his motivation and expertise. From this it would seem that those who are data savvy and express a personal interest in inquiry can also be looked to as “champions” of data and inspire social influence. Indeed, a person’s interest, commitment, and enthusiasm in evaluation is what Patton (2008) terms “the personal factor” and plays a major role in determining how much influence evaluation findings have. Many researchers have noted that social interaction and negotiation with colleagues are primary catalysts of data use within schools (Cousins & Leithwood, 1993; Simons, Kushner, Jones, & James, 2003; Spillane, 2012). It has been documented that the confluence of beliefs, knowledge, or motivations within groups has lead to shared understandings (Kennedy, 1982), the identification of different interpretations of the same data, and/or the construction of different actions in response to the data (Spillane, 2012).

Variations in individuals’ approaches to, and understanding of data challenge the Coburn and Turner (2011) representation of data use as a rational exercise. The notion that school practitioners methodically notice, then interpret, then construct implications from data connote a constructivist approach to information distillation and application. Alternatively, research in the domain of cognitive psychology suggests that individuals’ regard to data are not nearly so analytical. In particular, the contributions of Tversky and Kahneman (1975) investigate a number of mental operations, or heuristics, by which individuals exercise judgment in moments of decision-making. Their foundational work suggests that social biases lead to systematic errors in the ways in which individuals process information, naturally lending to errors in the “intuitive judgment of probability” in situations of uncertainty (p. 141). For example, in conducting a

series of experiments, Tversky and Kahneman (1973) determine that the “availability” of information – the plausibility of a scenario, or the ease with which a scenario comes to mind – can serve as the basis for a person’s judgement of the likelihood of a given outcome (p. 207). That, in fact, when decision-making moments are complex, people will tend to draw upon the simplest and most available scenarios in considering potential outcomes. As such, even though the “true” probability of an event is unknowable, individuals’ reliance on heuristics (like availability), are known to bias their subjective determinations of probability. Drawing on this seminal research, it is therefore recognized that the systematic integration of individual bias in weighing and synthesizing data in school-based contexts may complicate processes of data use in ways not fully captured by the Coburn and Turner framework (2011).

### **Organizational and Political Context**

As portrayed in Figure 2, school-based data use processes are embedded within an organizational and political context. The key dimensions of this context include “data use routines” that structure who educational practitioners interact with, around what data, and in what ways. Coburn and Turner (2011) emphasize that data routines encompass informal practices and highly-designed and structured activities, as well as naturally-occurring or evolving data activities. Their defining criteria for routines are “recurrent and patterned interaction that guides how people engage with each other and data in the course of their work” (p. 181). This concept of data use routines asks us to consider who exactly is involved in data conversations and the motivations, beliefs, and attitudes brought to the table in the consideration of data. Understanding the data use routine as a unit of analysis also requires insight into the specific type of data that are reviewed by schools (e.g., standardized test scores, student portfolios,

observations and experience), and how the attention of individuals is focused around that data. The concept of data routines thus provides a helpful contextual backdrop for the ways in which educators engage with one another in social interaction, in inquiry, and in approaching learning opportunities.

Coburn and Turner (2011) highlight a number of recurring factors observed in research involving data routines. The configuration of time (both the amount of time allowed for educators to collect, review, analyze, and interpret data, as well as the timeliness with which data is produced), and access to data (affected by technological infrastructure for housing and retrieving data, and the ways in which individuals are connected to each other within an organization), shape the creation of information. Organizational and occupational norms also guide the interaction of education practitioners. This is particularly salient in schools wherein norms of privacy have been seen to override interventions, encouraging teachers to talk specifically about their practice and share evidence of student learning with their colleagues. However, schools with norms that encourage teachers to share about their classroom practices openly, critique one another, or ask each other challenging questions have been seen to delve more deeply into issues of instruction and student learning (Little, 2007; Little, Gearhart, Curry, & Kafka, 2003a). Surveys of the factors specifically affecting the use of research in practice have been extensive, and Nutley and colleagues (2007) amalgamate these factors into four key areas: 1) the extent to which policy makers and practitioners are willing and able to use research; 2) the relevance of the research to practice; 3) the degrees of linkages between research and the policy and practice communities; and, 4) the context in which research use takes place are all considered fundamental to use.

*Leadership* is seen as an important factor across all of these dimensions, as are relations of power and authority. School and District leaders play a large part in selecting or designing data use routines, configuring time for teachers and others to engage in data use routines, deciding who gets access to what types of data, and establishing norms of interaction that involve trust and risk-taking and establishing data use as part of a school or district's culture (Feldman & Tung, 2001). The participation of school leaders in data use routines can steer conversations around the data – what is noticed and how it is interpreted – including the substance of the debate itself (Spillane, 2012).

*Power* and *authority* are extremely influential on data use routines as multiple stakeholders, each with different interests, pressure school administrators to pay attention to certain data and make certain decisions. In one direction, information can be used to reshape power dynamics between schools and their communities, such as for purposes of accountability. In the other direction, power dynamics have also been seen to influence data use, in particular what data individuals seek out and notice amid controversial issues or when backed by political motivations (Kennedy, 1982; King, 1988; Weiss, 1999). Levitt (2003) and Gabby and May (2004) suggest that the organizational and political context of data use in policy settings is, in fact, so dependent upon shifting power relations and agendas that data use behaviors often cannot be predicted.

As with their approach to understanding what cognitive processes individuals undergo in utilizing data, the Coburn and Turner (2011) portrayal of organization-wide decision-making “routines” references a constructivist tradition of organizational theory. Acknowledging this, the work of Scott (1981), and subsequently, Scott and Davis (2015), is instrumental in understanding how different theoretical perspectives, or paradigms of organizations may shape our

consideration of how “decision-making” is rooted within an institutional context. From a “rational system” perspective, Scott and Davis (2015) define organizations as “collectivities oriented to the pursuit of relatively specific goals and exhibiting relatively highly formalized social structures” (p. 29). This suggests that organizations maintain a rather distinctive character as well as a normative structure. From this viewpoint, organizations are oriented around specific goals that are translated into a set of prioritized preferences or functions which supply the criteria for choosing among alternative activities and which guide decisions about how an organizations structure is to be designed (p. 36). This is the analytic model underpinning the Coburn and Turner (2011) representation of schools as organizations.

Scott and Davis (2015) set forth two additional organizational paradigms under which schools may be alternatively considered. The first is the perspective of organizations as “natural systems” – one that focuses on the “behavioral structure” of the organization rather than its “normative structure.” Organizations as natural systems are defined as “collectivities whose participants are pursuing multiple interests, both disparate and common, but who recognize the value of perpetuating the organization as an important resource” (p. 30). From this viewpoint, what participants actually do rather than what they are supposed to do is a key element of consideration. How goals are implemented, as opposed to what is decided or planned, is the focus of “irrational” decision-making within a natural systems context. The second paradigm is that of the “open system,” wherein theorists view the organization as a system of interdependent activities and emphasize the multiple loyalties and identities of individuals comprising the organization. Open system organizations are defined as “congeries of interdependent flows and activities linking shifting coalitions of participants embedded in wider material-resource and institutional environments” (P. 31). This model stresses the importance of cultural-cognitive

elements in the composition of organizations. That is, organizations are observed to continuously adopt and adapt conceptions, models, schemas, and scripts, both intentionally and unintentionally, in their continuous production and reproduction of collective activity (p. 31). Decision-making takes place on a case-by-case basis and the ways in which decisions are made vary from individual-to-individual. Together, these three varying theoretical perspectives on organizations provide an important backdrop against which to consider the operational and relational aspects of schools as they experience decision-making processes and procedures. It remains probable that, although data are often presented as fuel for a rational approach to decision-making, this is only one way of understanding the value of data use from a certain organizational paradigm.

### **Interventions to Promote Data Use**

The complexity of data use processes must then be understood within the political and organizational contexts in which they take place. But these contexts are themselves subject to change. Coburn and Turner (2011) portray these influences as “interventions to promote data use,” the nature of which, they proclaim, shape the contexts and processes of data use in intentional and unintentional ways (p. 185). They summarize the wide variety of interventions introduced to promote data use in schools into three main categories: *tools*, *comprehensive initiatives*, and *accountability policy*. Tools as interventions include protocols for examining data, software systems that organize and create data reports (e.g., dashboards), new formative assessments, and/or processes for collecting and analyzing observational data. Comprehensive initiatives to foster data use are described as the incorporation of “multiple tools alongside professional development and new technology” (p. 186). Examples of this include school-based

or district-led inquiry projects focusing on wide ranges of school data, using protocols to guide data discussions, and involving trained facilitators or professional development. Lastly, accountability policies at the district, state, and federal levels have strongly promoted data use in schools. From this perspective, data is considered the main way to evaluate progress and is linked to incentives to change practice (Stecher, Hamilton, & Gonzalez, 2003).

Accountability policy is of particular interest within this study and warrants further review. While tools and comprehensive initiatives seem to be fairly targeted approaches to promoting data use, accountability policies are a much more indirect stimulus. The theory underlying accountability policies, as portrayed by Stecher, Hamilton, and Gonzalez (2003), is that student achievement will improve when educators are judged on student performance and when these judgments carry some consequences for educators. In focusing attention on student performance, schools create the need for increased data use and, subsequently, the practice of using findings from that data to encourage instructional change. Research on the effects of accountability policies on data use in schools, however, suggests that a wide variety of outcomes usually result. Several studies suggest that individuals regard the demand of accountability systems, as well as their responses to those systems, differently (Coburn & Turner, 2011). While accountability policies may be constructed to encourage particular data use behavior, there is great variation in the way individuals and organizations use data in response to these incentives. The work of Jennings (2012) attributes this differentiation to five fluctuating characteristics of accountability policies: 1) the expected pace for improvement on a continuum of supportive to punitive pressure; 2) the locus of pressure (e.g., districts, schools, teachers, or students); 3) the distributional goals set for students performance (e.g., growth vs. proficiency); 4) the features of assessments (e.g., content of student assessments); and, 5) the scope of the accountability system,

which may incorporate multiple measures or may be process- or outcome-oriented. King (1988) makes clear that the influence of accountability policies may, in fact, deter meaningful data use, drawing the distinction between data use for the purposes of signaling “compliance” with external accountability directives and data processes that provide practitioners with useful information for change.

Given these highly-modifiable elements of accountability policy, and the political and organizational contexts likely to play a part in defining them, it would seem that the relationship between data use interventions and the political-organizational realm is not unidirectional, as portrayed by Coburn and Turner. Coburn and Turner (2011) partially acknowledge this in their more detailed discussion of data use interventions that both interact with the political and organizational contexts, and which influence the process of data use: designed data use routines (e.g., teacher inquiry teams), technological tools (e.g., data dashboards), protocols and skilled facilitation, professional development, systems of meaning (i.e., categories, classification systems, and logics of action), and sanctions and rewards. In their discussion of sanction and rewards, Coburn and Turner (2011) mention the complexity of power dynamics characteristic of accountability systems that ensure schools and districts are responsive to their communities through the use of data, and which are increasingly being used as systems of monitoring and evaluation.

Perhaps an additional feature of data use interventions not fully addressed by the Coburn and Turner framework includes the guidance of experts. Research has shown that schools do benefit from the presence of experts who can assist teachers in management, reduction, analysis and interpretation of student data (Kerr, Marsh, Ikemoto, Darilek, & Barney, 2006), support teachers in applying the knowledge gained from student data to making instructional decisions

(Ikemoto & Marsh, 2007), and to assist teachers in identifying relevant research (Rock & Wilson, 2005). Experts may exist both “in-house” as instructional coaches or technical support for the use of data systems by teachers or principals, or as “external assistance” to improve the data output activities beyond school capacity (Gill et al., 2014), although some research warns that expertise is most effectively applied when support is provided by experienced, respected educators rather than technologists or statisticians (Datnow, Park, & Wohlstetter, 2007).

### **Potential Outcomes**

Lastly, Coburn and Turner (2011) address the intended outcomes of data use processes which are both wide in variety and span across multiple levels. From increased student learning to educators’ changed attitudes about student success to organizational learning, the use of data as a foundation for decision-making is perceived to have extensive potential. Coburn and Turner (2011) begin by discussing outcomes related to organizational change, which include changes in policy or strategic direction, changes in organizational structure, and changes in the way work and work roles are organized in education settings. Organizational change is considered at the school-level and district-level, as well as the system of public schooling writ large, but it is essentially portrayed as the sum of those changes in data use processes conducted by individual actors. This echoes earlier work by Lindblom and Woodhouse (1968) – who argued that what is politically feasible, in practice, involves only small-scale, incremental policy change – and Weiss (1982) – who termed the process of individual decisions that iteratively “coalesce and rigidify” into fixed results as “accretion.” While few education studies have sought to determine the impact of data use interventions on organizational change, Coburn and Turner (2011) suggest that of those that have, organizational change has been seen to result “when groups or individuals

engage in an iterative process of noticing, interpreting, and constructing implications for action in the context of data routines” (p. 192).

More research has been conducted to determine the effect of data use on changes in school administrator and teacher practices. In the case of district personnel, a change in practice might mean altering the ways they go about making a decision or implementing a policy, or the ways they work with each other and with schools. For principals this may mean new roles and responsibilities, as well as changes in the ways they interact with teachers, parents, and students. For teachers, changes in practice might entail altering instructional strategies, materials, and other classroom dimensions, as well as reshaping their roles within their schools and districts. Interventions introducing new data use routines for administrators have been shown, for example, to influence principals’ awareness, focus, and participation in the day-to-day academic plans and actions of teachers (McDougall, Saunders, & Goldenberg, 2007) or their propensity to restructure the school day to allow time for faculty dialogue and data-based inquiry (Feldman & Tung, 2001). Coburn and Turner (2011) emphasize that changes in administrator practice have important consequences for teacher practice, particularly because of administrators’ authority relations with teachers. Indeed, research has shown that when administrators strongly support data use routine, teachers tend to follow in their use of data to inform their teaching approaches (Feldman & Tung, 2001; Ikemoto & Marsh, 2007; Marsh et al., 2006). However, Coburn and Turner also recognize that this impact on teacher practices varies depending upon what school leaders emphasize and the nature of the data routines introduced.

Studies linking data use practices to student learning outcomes are few in number. Of those that have been conducted, changes in student learning outcomes seem to be the result of changes in teachers’ conversations about data via new data routines, protocols, and the active

participation of school administrators. As an example, Saunders and colleagues (2009) conducted a quasi-experimental, longitudinal study on the implementation of the “Getting Results” intervention. This introduced grade-level teams responsible for transforming academic standards into explicit instructional goals, identifying assessments and indicators to assess those goals, regularly evaluating school achievement and developing action plans, identifying and addressing teachers’ instructional challenges, aligning professional development with teachers’ needs, and facilitating regular grade-level team meetings focused explicitly on addressing identified student academic needs. Through an analysis of SAT-9 scores and state rankings, the authors found that the intervention produced significant school-level effects in comparison with controls (Saunders et al., 2009).

Improvements in student learning have also been detected as a result of teachers’ increased knowledge via the provision of professional development for data analysis. The work of Fuchs, et al. (1999) takes a close look at mathematics performance assessments which pose authentic problem-solving dilemmas and require students to use multiple skills and strategies to solve them. They found that teachers’ increased understanding of what the assessments are, as well as their knowledge of how such assessments could inform their instructional strategies, were linked with improved student learning gains, particularly in the case of above-grade students (Fuchs et al., 1999).

### **The Current Study**

The Coburn and Turner (2011) framework maps together current literature in order to provide a comprehensive portrait of data use in schools. As such, it also serves as a compass in guiding ongoing research. For example, it is apparent that while analyses of data use in specific school contexts contribute to an in-depth understanding of data use processes, it is just as

important to contextualize these actions from organizational and political perspectives. Although a challenging undertaking, there is a persistent, collective call from the research community for additional study of the effect of data use practices on organizational change, rather than solely the actions undertaken by individuals (Coburn & Turner, 2011, 2012; Nutley, 2007; Shulha & Cousins, 1997; Spillane, 2012).

As part of this, the Coburn and Turner (2011) framework presents a view to the improvement of school-based data use that is not yet tested in its application to schools in their day-to-day settings. Although Coburn and Turner (2011) acknowledge the dynamism of organizational and political contexts that interact and influence processes of data use, the overarching theory of action introduced by the framework suggests that certain stimuli (i.e., tools, comprehensive data initiatives, and policies) will positively influence schools' use of data and culminate in organizational change, changes in practice, and/or student learning. That is, the effective use of data in school-based decision-making is an outcome of establishing the correct procedures or introducing the appropriate resources, which will in turn improve the state of teaching and learning within a school. In a similar vein, the Coburn and Turner (2011) approach assumes that stakeholders also engage in rational processes of data filtration and synthesis (i.e., noticing data, interpreting data, and constructing implications for data). This implies a naturally systematic approach to prioritizing and applying data in logical processes of decision-making undertaken at the institutional level.

This study returns to the idea that data used for decision-making may occur in non-rational, non-linear processes and that the effective use of data may be more reliant upon the orientation of a school's organizational culture rather than the simple introduction of inputs. In exploring how data are identified by different school stakeholders, the ways in which they are

valued, and the ways they are incorporated into processes of decision-making, this study addresses processes of meaning-making occurring at both the individual and organizational levels. That is, while it is understood there exist a host of possible types of “data” under Ackoff’s broad definition, and a variety of ways in which data might be “used,” this study is focused on examples of data raised by school practitioners as they consider what sources of information are integrated within their day-to-day work. The perceived value of those data, practitioners’ expectations around how data ought to be used, as well as descriptions of whether and how data are eventually applied substantiate this study’s inquiry around how data are, or are not used to inform decisions around instruction and student and school performance. In so doing, this study attempts to uncover how individuals make sense of the various types of data available within schools and how these diverse perspectives influence school responses to data demands, including those imposed by accountability and self-evaluation activities. It further strives to depict how data are perceived and used in the day-to-day functioning of schools, and how relationships between organizational and individual efforts to use data are negotiated and established. Ultimately, in the acknowledgement of school-based approaches to data identification, interpretation, and use, this study sets aside the Coburn and Turner (2011) framework and using a grounded approach, explores the ways in which the actual work of schools and individual practitioners incorporates, ignores, understands, and evaluates data within their experiences of decision-making.

## **CHAPTER 3**

### **RESEARCH METHODS**

#### **Introduction**

This section details the methods and analyses employed in understanding the use of data in school-based contexts for purposes of decision-making in response to the question, *how do teachers, principals, and district personnel use data in their professional contexts?* More specifically:

1. What do school practitioners identify as data, and particularly as credible data?
2. How do teachers and principals use data to inform decisions related to school improvement and strategic planning?
3. How do teachers use data to inform instruction?
4. How do teachers, principals and district personnel use data to monitor school performance?
5. How do organizational and cultural characteristics of schools affect the way teachers and principals use data (for any of those purposes)?

#### **Study Procedures**

A cross-case comparative approach was applied to this study in an attempt to qualitatively investigate interpretations of data use practices in schools. To follow is a description of how this method was employed as a way of providing specific description of the processes and contexts of data identification, interpretation, and use influencing individuals within three high schools within the Los Angeles Unified School District. Also included are procedures applied in the conduct of an online teacher survey aimed at understanding more general patterns in perspectives on, and experiences with data use in a quantitative dimension. Due to low response rates, however, these results are not discussed in the study's final findings.

## Study Setting – Pilot Schools

The effective use of data to inform decision-making has become acknowledged as a key best practice in well-performing schools; but, as the preceding literature review shows, there is little research offering a window into the ways in which teachers, principals, district administrators, parents, and others identify, interpret, and use data for decision-making in ways that are systemically effective. The LAUSD pilot schools, however, present a particularly intriguing context in which to examine individual-, organization-, and system-level processes of data-driven decision-making.

Established in 2007, *pilot schools* are a network of public schools granted charter school-like autonomy over six key areas: budget, curriculum and assessment, governance, professional development, school calendar and scheduling, and staffing (Martinez & Quartz, 2012). Created to be models of education innovation, pilot schools feature professional learning communities and a unifying mission and vision, are small in size (optimally 400-500 students), are self-governed and led, and are expected to be research-based, student-centered, and strong partners with parents and their communities (“LAUSD Pilot Schools,” n.d.). As part of this arrangement, pilot school teachers remain members of the United Teachers Los Angeles (UTLA) union but operate under a “thin contract” which allows teachers to work extra hours. Despite these overarching characteristics, individual pilot schools are unique in their exercise of the various autonomies such that each campus implements its own, tailored strategy for sustained improvement.

In exchange for their greater organizational autonomy, pilot schools are subsequently subject to strong accountability measures. Each pilot school is expected to conduct annual self-reviews and longitudinal data monitoring, and to field scheduled visitations by external review

teams. The District states that the goals for these activities are to “initiate meaningful dialogue among school stakeholders, provide substantive feedback on strengths, challenges, and recommendations for improvement, to assess school progress across multiple indicators of student engagement and achievement, and to provide data to key stakeholders” (Los Angeles Unified School District, 2012). Given this focus on both autonomy and accountability, the pilot school initiative presents intriguing questions with respect to how individuals within schools, schools as organizations, and pilot schools as a collective define, interpret, and make use of evidence to inform their own progress and performance. Because pilot schools are expected to develop their own individual theories of change, the ways in which each school evidences the success of its strategic vision become interesting points of comparison and contrast in understanding schools’ effective use of data in decision-making (Small, 2009). As compared with their conventional high school counterparts, which must respond to mandated data-based activities and District requests, the pilot schools serve as examples for schools’ *potential use* of data in decision-making moments and present a possible “best case scenario” in which data, and data use activities may be flexibly exercised.

### **Comparative Case Study**

A comparative case study was conducted among three pilot high schools in LAUSD as a way of investigating key aspects of data identification and use, as well as similarities and differences across pilot high school sites (Anckar, 2007; Liphart, 1975; Ragin, 1994; Yin, 2003). The case study approach addresses the embedded nature of school knowledge systems (e.g., individual, organization, and system) and the many data use factors that play out at each of these levels (Scholz & Tietje, 2002; Yin, 2003). Additionally, this approach can incorporate variables

of interest exceeding the number of cases feasibly included within the study (Campbell, 1975; Mahoney, 2000). Case studies are particularly useful wherein there is little control over behaviors of interest, and take a holistic approach to the complexity of school systems by relying on multiple sources of evidence (i.e., interviews with multiple stakeholders, participant observation, and document review) through which a convergence of findings was sought.

### Case Study Design

In 2013-2014, LAUSD had officially established Memorandums of Understanding with 48 pilot schools. Of these, 36 schools taught Grades 9-12 and 4 schools were “span schools” offering either Grades 6-12 or K-12 instruction. To ensure a sufficient number of participant candidates, as well as a reasonable degree of comparability between cases, only pilot school sites offering programming for Grades 9-12 were considered for this case study participation. Three (3) pilot high schools were purposively selected as the comparative case study sample.

Of primary interest to the study is the comparison and contrast of schools in their approach to, and perceived success in, conducting self-evaluation activities. School leadership is known to be a key element in the success of school-based data collection and use (Feldman & Tung, 2001; Ikemoto & Marsh, 2007; Marsh et al., 2006). Minimal criteria for case selection thus required that school principals had some working knowledge of data use activities occurring within their school, either in response to accountability requirements or with respect to school-developed data use initiatives. While it was initially perceived that the final sample should include pilot high schools representing “emerging,” “middle-of-the-road,” and “highly-successful” evaluation systems, these classifications were ultimately found to be inadequate in application. A meeting convened with pilot school district managers revealed that the development of pilot school data use systems had not been regarded in this way, and consensus

around these criteria was not definitive. Each schools' use of data was found to be so contextually nuanced that the comparison of school "evaluation systems" *ipso facto* was inappropriate. Alternatively, it was found that the number of years each pilot high school had been active was strongly associated with the maturity of its systems and processes of data use. As a result, schools representing different years of operation were selected for participation. Case study enrollment also depended on a schools' willingness to accommodate multiple interviews with the principal and at least three teachers, as well as participant observation.

Pilot high school sites were initially recruited via a formal letter sent to principals from LAUSD's Superintendent's Intensive Support and Intervention Center (ISIC), yielding 5 volunteers. An additional 8 pilot high school principals were contacted by phone following the collection of background information and candidate suggestions from multiple sources, including: LAUSD personnel within ISIC; a local education non-profit working with several pilot high schools and its external evaluator, a founder of the pilot school initiative within LAUSD; and, a pilot high school principal.

Based on the criteria outlined above, the final sample of case study participants include one of LAUSD's first established pilot high schools in its sixth year of operation, a four-year-old pilot high school, and a recently-established pilot high school in its second year of operation. School sites were not selected as final study participants if they did not yet offer complete programming for Grades 9-12 or if they were unable to host multiple interviews with the principal and at least three teachers. One principal of a fourth pilot high school was able to offer her time for two interviews and recommend at least one of her lead teachers for a single interview. Data collected from this fourth school are referenced within the study but do not substantiate a complete school case. Additional sample details are outlined in Table 1.

*Table 1: Case Study School Participant Characteristics*

| School Name*                                       | Year Opened | Years of Operation at Time of Data Collection | Grades of Instruction | Notable School Characteristics                                                                       |
|----------------------------------------------------|-------------|-----------------------------------------------|-----------------------|------------------------------------------------------------------------------------------------------|
| The Academy (Case #1)                              | 2013        | 2                                             | 9-12                  | Co-located on campus with conventional high school.                                                  |
| Belleworth School of Arts and Technology (Case #2) | 2011        | 4                                             | 9-12                  | Co-located on campus with several other pilot schools. First year with new principal.                |
| Woodson College Preparatory School (Case #3)       | 2009        | 6                                             | K-12                  | Co-located on campus with several other pilot schools. Partnered with research intensive university. |

\*Pseudonyms have been assigned for the protection of participant identity.

### Participant Selection and Data Collection

Within each school site, principals were asked to recommend teacher study candidates representing varying degrees of interaction with school data and evaluation practices. Teachers were contacted by email and phone, and in some cases were approached in person following an introduction to faculty by the principal. Criteria for teacher selection included willingness and availability to participate in interviews at least three times over the course of the year, as well as some knowledge or understanding of school data and evaluation practices. Teacher samples were composed to include at least one individual with intimate knowledge of these activities (e.g., someone who led or coordinated evaluation or assessment activities), although all individuals had at least general day-to-day experience with data and data use processes. Teacher volunteers in excess of the minimum three were also enrolled within the study provided that they expressed some understanding of data use activities within the school. Staff members (i.e., out-of-class faculty) extensively involved in school data use activities and initiatives were also approached for study participation when recommended by the principal. While teacher participants were not

selected based on their numbers of years of experience or their status as lead teachers, these characteristics were observed to influence participants' perspectives as data users and are provided as descriptive background. The final sample of principal and teacher interviewees is presented in Table 2 below.

*Table 2: Case Study Teacher and Principal Participant Characteristics*

| School Name*                                                      | Principal Name* | Teacher Name* | Lead Teacher? | Teaching Faculty? | Number of Years Teaching |
|-------------------------------------------------------------------|-----------------|---------------|---------------|-------------------|--------------------------|
| <b>The Academy<br/>(Case #1)</b>                                  | Mr. Cooper      | --            | N/A           | Yes               | 27**                     |
|                                                                   | --              | Mr. Leighton  | N/A           | Yes               | 19                       |
|                                                                   | --              | Mr. Easton    | N/A           | Yes               | 27                       |
|                                                                   | --              | Ms. Hanley    | N/A           | Yes               | 19                       |
|                                                                   | --              | Mr. Knowles   | N/A           | Yes               | 8                        |
| <b>Belleworth School of<br/>Arts and Technology<br/>(Case #2)</b> | Ms. Heredia     | --            | N/A           | No                | 13**                     |
|                                                                   | --              | Ms. Gavin     | Yes           | Yes               | 5                        |
|                                                                   | --              | Ms. Nava      | Yes           | Yes               | 5                        |
|                                                                   | --              | Mr. Nuñez     | No            | Yes               | 5                        |
|                                                                   | --              | Ms. Salceda   | Yes           | Yes               | 5                        |
| <b>Woodson College<br/>Preparatory School<br/>(Case #3)</b>       | --              | Mr. Neal      | No            | Yes               | 9                        |
|                                                                   | Ms. Figueroa    | --            | N/A           | No                | 18**                     |
|                                                                   | --              | Mr. Macon     | Yes           | Yes               | 16                       |
|                                                                   | --              | Ms. Lovell    | Yes           | Yes               | 9                        |
|                                                                   | --              | Mr. Urbina    | Yes           | Yes               | 11                       |
|                                                                   | --              | Ms. Gilman    | No            | Yes               | 13                       |
|                                                                   | --              | Dr. Baher     | No            | No                | N/A                      |
| Foxvalley School of Arts<br>and Music<br>(Supplementary Data)     | --              | Ms. Finche    | No            | No                | 11                       |
|                                                                   | Ms. Davila      | --            | N/A           | No                | ***                      |
|                                                                   | --              | Ms. Lam       | Yes           | No                | ***                      |
|                                                                   |                 | Ms. Owen      | No            | Yes               | ***                      |

\*Pseudonyms have been assigned for the protection of participant identity.

\*\*Inclusive of years teaching and school administration.

\*\*\*Data not available.

Principals and teachers at each school site participated in multiple interviews throughout the academic year in order to gain their perceptions on school-based evaluations, data identification, and data use processes. On average, principal interviews were each about one hour in length, while teacher interviews were about 40 minutes. Interviews of staff members or

teachers who could only participate in one interview were extended to about 1.5 hours. All interviews were semi-structured, adhering to a general set of topics and themes outlined in the interview protocol. Questions for teachers and principals revolved around understanding school and teaching performance objectives, perceptions of “information,” understanding school accountability requirements, perceptions of data use, school culture, technical capacity, and data use policies and tools. Study design and interview protocols were reviewed and approved by the UCLA Institutional Review Board (UCLA IRB#: 14-000849).

Additionally, observations of professional development meetings, committee meetings (e.g., Governing School Council meetings), and meetings convened around specific data initiatives (e.g., student assessment) were conducted at all three school sites following participant invitation. These observation periods were, on average, one hour in length. Intensive observation and participant observation conducted during school-based data collection and review activities, however, spanned one to two days. Documents collected from observations (and, in some cases, interviews) included meeting agendas, copies of presentation content, photographs of school campuses, memos, and reports. The complete schedule of interviews and observations throughout Academic Year 2014-15 is presented in Table 3 below. This timeline reflects the early recruitment of The Academy into the study, with later participation by Belleworth School of Arts and Technology and Woodson College Preparatory School. Staggered study enrollment was a result of school site availability. Likewise, the timing of participant interviews was subject to principal, teacher, and staff availability.

Table 3: Interview and Observation Details

| School Name                                                      | Participant Name    | Academic Year 2014-15 |     |     |     |     |     |     |     |     |     |     |     | Total Interviews & Observations |
|------------------------------------------------------------------|---------------------|-----------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|---------------------------------|
|                                                                  |                     | Sep                   | Oct | Nov | Dec | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug |                                 |
| The Academy<br>(Case #1)                                         | <b>Interviews</b>   |                       |     |     |     |     |     |     |     |     |     |     |     |                                 |
|                                                                  | Mr. Cooper          | x                     |     | x   |     | x   |     | x   |     |     | x   |     |     | 5                               |
|                                                                  | Mr. Leighton        |                       |     | x   |     | x   |     | x   |     | x   |     |     |     | 4                               |
|                                                                  | Mr. Easton          |                       |     |     |     |     |     |     | x   |     |     |     |     | 1                               |
|                                                                  | Ms. Hanley          |                       |     |     | x   |     | x   |     | x   |     |     |     |     | 3                               |
|                                                                  | Mr. Knowles         |                       | x   |     | x   |     |     | x   |     |     | x   |     |     | 4                               |
|                                                                  | <b>Observations</b> |                       | x   | x   |     |     |     |     |     |     | x   |     |     | 5                               |
| Belleworth School<br>of Arts and<br>Technology<br>(Case #2)      | <b>Interviews</b>   |                       |     |     |     |     |     |     |     |     |     |     |     |                                 |
|                                                                  | Ms. Heredia         |                       |     |     |     | x   |     | x   |     | x   |     |     |     | 3                               |
|                                                                  | Ms. Gavin           |                       |     |     |     | x   |     | x   | x   |     |     |     |     | 3                               |
|                                                                  | Ms. Nava            |                       |     |     |     | x   |     | x   |     | x   |     |     |     | 3                               |
|                                                                  | Mr. Nuñez           |                       |     |     |     |     |     | x   | x   | x   |     |     |     | 3                               |
|                                                                  | Ms. Salceda         |                       |     |     |     |     |     | x   | x   | x   |     |     |     | 3                               |
|                                                                  | Mr. Neal            |                       |     |     |     |     |     | x   |     | x   |     |     |     | 3                               |
|                                                                  | <b>Observations</b> |                       |     |     |     |     | x   | x   |     |     |     |     |     | 6                               |
| Woodson College<br>Preparatory<br>School<br>(Case #3)            | <b>Interviews</b>   |                       |     |     |     |     |     |     |     |     |     |     |     |                                 |
|                                                                  | Ms. Figueroa        |                       |     |     |     | x   |     |     |     | x   | x   |     |     | 3                               |
|                                                                  | Mr. Macon           |                       |     |     |     |     |     | x   | x   | x   |     |     |     | 3                               |
|                                                                  | Ms. Lovell          |                       |     |     |     |     |     | x   | x   | x   |     |     |     | 3                               |
|                                                                  | Mr. Urbina          |                       |     |     |     |     |     | x   | x   |     | x   |     |     | 3                               |
|                                                                  | Ms. Gilman          |                       |     |     |     |     |     | x   | x   | x   |     |     |     | 3                               |
|                                                                  | Dr. Baher           |                       |     |     |     |     |     | x   | x   |     | x   |     |     | 3                               |
|                                                                  | Ms. Finche          |                       |     |     |     |     |     |     |     | x   |     |     |     | 1                               |
|                                                                  | <b>Observations</b> |                       |     |     |     |     |     | x   | x   | x   |     |     |     | 8                               |
| Foxvalley School<br>of Arts and Music<br>(Supplementary<br>Data) | <b>Interviews</b>   |                       |     |     |     |     |     |     |     |     |     |     |     |                                 |
|                                                                  | Ms. Davila          |                       |     | x   |     | x   |     |     |     |     |     |     |     | 2                               |
|                                                                  | Ms. Lam             |                       |     |     | x   |     |     |     |     |     |     |     |     | 1                               |
|                                                                  | Ms. Owen            |                       |     |     | x   |     |     |     |     |     |     |     |     | 1                               |
|                                                                  | <b>Observations</b> |                       |     |     |     |     |     |     |     |     |     |     |     | 0                               |

Finally, interviews were held with two LAUSD Intensive Support and Intervention Center (ISIC) personnel, alongside observations of a District-facilitated pilot school conference and a Pilot School Steering Committee meeting. Interviews were also semi-structured and included questions pertaining to the development and management of pilot schools; policies, guidelines, and processes applied in measuring pilot school performance; and perceptions of what constituted “successful” school performance. District-level interviews and observations are detailed in Tables 4 and 5 below.

*Table 4: District Interview Details*

| <b>Participant Name*</b> | <b>Title</b>                                   | <b>Total # of Interviews</b> | <b>Date</b>                      | <b>Avg. Interview Length</b> |
|--------------------------|------------------------------------------------|------------------------------|----------------------------------|------------------------------|
| Ms. Macia                | ISIC - Director of Autonomy and Accountability | 3                            | Sep 2014<br>Mar 2015<br>Jul 2015 | 1-Hour                       |
| Ms. Noriega              | ISIC - Instructional Director                  | 1                            | Sep 2014                         | 1-Hour                       |

\*Pseudonyms have been assigned for the protection of participant identity.

*Table 5: District Observation Details*

| <b>Activity</b>                 | <b>Date</b> | <b>Observation Length</b> |
|---------------------------------|-------------|---------------------------|
| Pilot School Conference         | Sep 2014    | 6-Hours                   |
| Pilot School Steering Committee | Jun 2015    | 2-Hours                   |

It should be noted that the original study design called for interviews with at least two parents per school site. Candidates were contacted via each school’s community outreach coordinator, community coordinating center, or principal. Parent availability, however, was limited and recruitment efforts resulted in interviews with 2 parents from The Academy, 2 parents in a joint interview at Belleworth School of Arts and Technology, and 1 parent at

Woodson College Prep. While parental perspectives are considered both relevant and important to the topic of data use in schools, because parent data were relatively sparse across schools, and in some places unreliable (there were substantial issues with language translation at Belleworth School of Arts and Technology), these data have not been included within this study. However, they may be reserved for future research regarding parental interaction with data use practices in schools.

### Analytic Procedures

Each selected school site represents a unique case within this study, and within which principal and teacher participant are embedded (Yin, 2012). As such, data were collected within each school and analyzed using Erickson's (1986) approach to interpretive analysis whereby induction and deduction are in constant dialogue. Initial analysis was conducted without reference to previous theoretical propositions as a way of “playing with the data” in their unique contexts and in search of naturally emerging patterns or concepts. A second stage of analysis drew on the current Coburn and Turner (2011) conceptual framework describing data use in schools. The process of analysis was continuous, and findings were constructed as pieces of data gathered to both reflect specific lines of inquiry and in the need to adapt those lines of inquiry to contextual events in school settings. Data were separately analyzed for each school case. Once complete, more general abstractions across cases were constructed in a cross-case analysis (Yin, 2012). The ultimate objective was to produce findings in the form of organized descriptive accounts, themes, or categories that cut across the data, contributing to a conceptual model that explains the data (Merriam, 1998).

To start, all interview recordings were transcribed and combined with observation field notes in order to derive a specific understanding of meaning-making through the documentation of concrete details of practice. Data was imported into, and analyzed using, NVivo software, wherein a process of open coding was applied to identify segments of data that might be helpful in answering the research questions. Next, a process of axial coding was applied to search for commonalities and assertions that “hang together” across participant data (Saldaña, 2015). A final list of 228 codes was derived from the data and is presented in Appendix A. This process drew on interpretation and searched for comparative understandings of local meanings, social settings, and constructs beyond the immediate circumstances of local settings (Erickson, 1986; Richards, 2009). The internal validity of results was tested through the triangulation of participant data and document review. Additionally, analyses were also subjected to “member checks” by key interview participants, who reviewed content for accuracy of interpretation. Rich, thick description ensures that sufficient data is provided for transferability and potential extrapolations, and continuous reviews of analyses ensure that results are consistent with the data collected (Lincoln, 1985; Patton, 2005). Disconfirming evidence was actively sought as a way of identifying alternative ways of presenting the data or contrary explanations (Merriam, 1998).

### **Pilot High School Teacher Survey**

In an effort to determine whether the perceptions and attitudes toward data use in schools uncovered in the cross-case comparison might be representative of those held by a more general population of pilot high school teachers, an online survey was distributed to pilot high school teachers throughout LAUSD. Survey content was based on preliminary analyses of interview transcripts, interview notes, and observation notes. The final survey was designed to be

completed in 10-15 minutes and included 25 items requesting teachers' background information asking them to identify "useful" sources of information for various purposes, and soliciting comments on their beliefs, experiences, and perspectives around the use of data in their school. An initial draft of the survey was piloted with three teacher participants from the case study, subsequently revised, approved by UCLA's Institutional Review Board, submitted to the Office of Data and Accountability for review and approval, and administered online via SurveyMonkey on September 1, 2015.

Unfortunately, in accordance with LAUSD's privacy policies, direct contact could not be made with individual teachers via email or phone to solicit participation. Principal contact information was obtained through LAUSD's public directory and principals were requested by email in early August 2015 to forward survey invitations to faculty teaching Grades 9-12. Individual phone calls to principals were made alongside the email invitation the same week. Two additional follow-up emails were sent to principals on September 1 and September 15, 2015. The survey closed after three weeks on September 22, 2015. In total, 93 teachers across 13 pilot high schools and span schools consented to participate in the survey; of these, 87 indicated they were classroom teachers (the remaining 5 were administrators or counselors and could not be verified as classroom teachers). This represents about one-third of all pilot schools offering Grades 9-12 and a 7% response rate among all pilot school teachers with classroom rosters for Grades 9-12. Due to this low response rate, statistical power was not sufficient enough to treat survey results analytically. As a result, these data are not presented within this study's findings. However, the survey instrument will be retained for application to future research.

## CHAPTER 4

# SCHOOL DATA SYSTEMS AND STRUCTURES

### Introduction

Each of the three school cases within this study presents a unique orientation to the use of data in school-based decision-making. Identified through observations, interviews, and surveys, four general categories of culturally-defined structures and practices appeared to shape each school's relationship with school-based data: 1) how each site determines who is charged with making what decisions, 2) how those decisions are made, 3) how school stakeholders conclude what constitutes "credible data," and 4) how processes of data use are established and developed.

Underlying even these fundamental contextual factors, however, are the basic systems and structures intentionally constructed by each school to initiate, support, and refine data use within everyday school activity. These include policies and procedures by which data are collected, analyzed, and reported, such as scheduled time for teachers to review, deliberate, and collect data, as well as plan and design assessments and evaluations. Previous findings, such as those presented by Coburn, et al. (2009), suggest that where adequate time is not provided within schools to debate conflicting interpretations of data, and to evaluate different solutions, decision-making can be *conservative*, *prolonged*, or altogether *unresolved*. Certainly the technological infrastructure underlying data access within schools has been observed as a key factor in data collection, storage, and retrieval (Lachat & Smith, 2005; Marsh et al., 2006; Means, Padilla, DeBarger, & Bakia, 2009; Means, Padilla, & Gallagher, 2010; Thorn, 2001; Wayman, 2007; Wayman, Conoly, Gasko, & Stringfield, 2008; Wayman, Stringfield, & Yakimowski, 2004). Human infrastructure is also acknowledged for the ways in which it influences how individuals

in different parts of an organization are connected to one another, subsequently impacting flows of information (Coburn, 2010; Daly & Finnigan, 2011; Honig, 2006).

Criteria for participation in this study required that each school participant express some degree of understanding or commitment to the explicit use of data in making school-based decisions. However, the extent to which each school had proactively developed guidelines, policies, and systematic data use processes varied considerably. These formal structures are viewed as important building blocks to other organizational and political dimensions – they are the backdrop against which organizational expectations guide regular data use. Throughout the course of this study, it has become apparent that the strength of each school’s data use systems and structures are strongly tied to its maturity and development as a pilot school. At the time of data collection in academic year 2014-15, the first-established pilot schools were, at most, eight-years-old. Also at this time, the LAUSD was working to expand the pilot school model and many sites had only just been formed. As a result, LAUSD pilot schools vary not only in their formal approach to data use, but also in the development of their governing systems and infrastructure.

Take for example The Academy which first opened its doors in 2013. By 2014, it was still considering what systems and guidelines needed to be in place to establish regular reviews of data. Belleworth School of Arts and Technology, while established in 2011, experienced a recent change in principal leadership in 2014. With this administrative shift came a complete reorientation to the use of data and their contribution to improvements in teaching and learning practices. Lastly, Woodson College Prep was one of the first pilot schools to be approved by the District, and in its sixth-year of operation, has formed a rather robust program of data use activities, one that it consistently continues to refine.

Given the confluence of data use procedures with school operational systems, the three school cases presented throughout this study cannot simply be regarded as representations of three incremental “levels” of data use proficiency. Rather, each school must be understood in the context of its current state of development and the unique aspects of its organizational environment. Summary background for each of the school sites is presented within this chapter as a way of understanding the circumstances undergirding its use of data in decision-making.

### **Case #1: The Academy**

George Washington High School (GWHS) is a large comprehensive school within LAUSD. It spent several years implementing a Small Learning Community (SLC) model on its campus wherein it was anticipated that students would benefit from participation within small, distinct learning groups driven by a focus in content interest, such as health care, human services, or visual and performing arts. Despite its rather robust SLC programming, for a variety of reasons GWHS drastically minimized the visual and performing arts-focused SLC. The teachers comprising that small learning community considered applying to become a separate, self-governed pilot school. For three years, a design team consisting of about ten GWHS teachers developed their pilot school proposal. They selected a principal, Mr. Cooper, from outside LAUSD as someone who brought in 27 years of teaching experience in both a well-known high school and a well-reputed college-level teaching credential program. This would be Mr. Cooper’s first position as a school administrator. In the 2012-13 academic year, with a majority staff vote and District approval of their proposal, the visual and performing arts SLC broke away from GWHS as an independent pilot school called The Academy.

In its first year of operation, the 2013-14 academic year, The Academy's student body included just under 400 students, the majority of whom were Latino, about 25% being African American, about 10% being White, and the remainder being Filipino, Pacific Islander, Asian, and American Indian, or Alaska Native. Over 70% of its learners were considered socioeconomically disadvantaged, and less than 10% were categorized as "English language learners" (ELLs). Just over 5% of the student body represented students with disabilities. In its second year of operation, 2014-15, The Academy enrolled a slightly higher number of students, with a total enrollment still hovering around 400, and maintained a similar student demographic representation.<sup>2</sup> Enrollment is anticipated to continue increasing, a sign of The Academy's health and growing traction within the community.

The Academy's independence as a pilot school, however, did not transpire into a geo-physical break from GWHS. With no District plans to build additional school sites, the only school site available to host The Academy as a campus was within its "parent" school. GWHS was subsequently mandated by the District to turn over some of its buildings and administration offices to the new pilot school. Perhaps as no surprise, an antagonistic relationship developed between GWHS and The Academy, impacting The Academy's ability to actualize its own culture and autonomy. As one Instructional Director from the District put it:

*How do you function at a school where not only are you saying, "Well, we're a new school, like it or not, we're here," to, "Not only are we a new school, but we're sharing this campus with someone, who has been there... you know a school that has been here for 60, 70, 80 years, and now you've got two principals? Two schools?"*

---

<sup>2</sup> School profile data for AY 2014-15 was not publicly available from LAUSD at the time of writing.

This sour relationship manifested itself not only as a challenge to The Academy's individual culture, but also to its basic operational structure. The Academy's initial enrollment of just under 400 students in 2013 was far lower than predicted. The teachers and principal of The Academy credit the principal of GWHS for actively suppressing the number of incoming students by, for example, withholding GWHS students who would otherwise have transferred and dissuading others with news that The Academy was not offering AP courses. With a low student-to-teacher ratio unsustainable by the District, several teachers from The Academy – including members of the original design team – faced mandatory dismissal just after the academic year had begun and enrollment figures were finalized.

This sudden decrease in staff served as a blow to the sense of confidence and security of The Academy's new faculty, as well as to its pool of human capital. As one teacher described it:

*That first year, we lost all of these teachers, and everybody was on the... kind of like, in defense mode and survival mode.*

“Survival mode” for this particular teacher meant having to significantly step up his own level of effort both in terms of teaching as well as in actively administering the school.

*My own personal situation is that I helped start the school, I... and whatever happens is the kind of thing... I'm going to work on that. So I've got my hands in everything, which means... we had to collapse at the beginning of the year because of our numbers. We had to collapse an English line, so I ended up instead of teaching just Photo and having a conference period, I teach seven periods and have no conference period and am teaching four preps.... And then doing all the other stuff to help make the school work.*

The greater challenge underlying this unexpected depletion in workforce, however, is the simple fact that the school was launching into its inaugural year. In the words of another seasoned teacher who had been recruited into The Academy's faculty:

*Just being a brand-new school, we have so many other issues to also deal with and tackle. At some schools it would just be routine, because those schools have been around 30, 40, 50, or 60 years. And certain things are just done on autopilot. As a school that's only been open for a year and a half, NOTHING is on autopilot. And, as a result... we've got decisions to make about everything from graduation and senior pictures to... any other myriad of things that kind of come up.*

Laying down the groundwork of The Academy – all of its systems, policies, processes, and procedures – had been no easy feat. The motivation to bring its concept to life and to institutionalize itself as a model of good practice reverberates in the school's tenor and pace. However, the growing pains experienced in establishing The Academy had been real and urgent. In these early days, ensuring that The Academy is functionally operating and meeting all legal school requirements often takes precedent as “emergency” demands. Given the immediacy and volume of these new school needs, it is understandable that formal systems of data collection, analysis, interpretation, and use had not yet been established. As expressed by another Academy teacher, there simply was no time with which to systematically review school data. Even figuring out how to access student data through the District's new information system had been an uphill and arduous battle.

The Academy was not alone in its startup frenzy. Indeed, the two other pilot high schools included within this study highlighted similar feelings of operational stress and chaos in their own initiation. The Academy thus presents one example of how a school determines what data is collected, for what purposes, and how they are used and valued when no formal systems of data flow exist. It provides some insight into how a new school thinks about how well it is doing while determining its benchmarks of success. That is, when everything seems to be in the flux of

development, what do teachers and the principal gravitate toward to understand student and school performance?

## **Case #2: Belleworth School of Arts and Technology**

Displaced from their positions at Evenwright High School,<sup>3</sup> a team of teachers gathered in 2010 to form the design team for what would eventually become Belleworth School of Arts and Technology. While the District had approved the teacher group to function as a design team for the new pilot school, following the ratification of their proposal, it was ultimately determined that these teachers could not be guaranteed a teaching position within the very school they worked to design.

When Belleworth opened its doors in 2011 on a large campus co-located by three additional pilot high schools, the implementation of Belleworth's original mission and vision was left in the hands of an entirely new administration and faculty. None of the teachers from the original design team was on faculty (this was true for all the campus's pilot school sites), and despite the design team's interest in self-selecting Belleworth's first principal, this position was filled by the District. Although teachers from Evenwright were given priority to assume positions at Belleworth (based on their seniority and standing with the District), none applied. Belleworth thus began with an entirely fresh batch of administration and teachers divorced from both Evenwright and from Belleworth's original pilot proposal.

The establishment of Belleworth's basic operational systems was further complicated by year-to-year changes in its administrative structure. As depicted by one teacher participant, Ms.

---

<sup>3</sup> This school had undergone a formal process of "reconstruction" under the purview of LAUSD. Under this particular reform measure, LAUSD mandated a district takeover of Evenwright, the authority for which was granted under the No Child Left Behind Act of 2001.

Gavin, the District initially maintained a heavy hand in how Belleworth and its three other co-located pilot schools would work together. She described that in the first year the campus was opened, all four schools were essentially treated as a single, comprehensive school. She recounted, for example, how students were constantly shuffled between schools, such that she had an entirely different set of students after one month of teaching, again at the end of the semester, and again following the winter break.

Ms. Gavin recalled Belleworth's second year allowed limited exchanges of students between the pilot schools, and school organization was "a little better." She continued:

*Third year... finally, it was autonomy. Separate schools, separate disciplines, and our third year, which was last year, it was probably the first year where schools actually – and our school specifically – got to feel like we were going to start working on our mission and our vision.*

Alongside the configuration of how the pilot schools would be organized as a collective, the leadership style brought by Belleworth's first principal allowed for a great deal of flexibility in how the school would be governed. Belleworth's incoming faculty, therefore, was in many ways responsible for administrative duties as well as instruction, curriculum development, and establishing the mission and vision of the school. Ms. Gavin remarked:

*Our other leader was really like, go ahead, take it on yourselves. And that was difficult because we didn't know... all the resources we had, we didn't know... what we were supposed to be doing.*

From her perspective, this "hands-off approach" entailed a substantial degree of work and associated stress among the staff, none of who had been trained in school administration.

In fact, she emphasized that it wasn't until last year that Belleworth was able to start working on its individual identity as a pilot school. And with the retirement of its first principal, another change in leadership in 2014 – a hiring process this time led by Belleworth faculty – was

looked to as an opportunity to find someone who could both support and guide staff in the continued development of the school. Ms. Gavin reported that “it wasn’t until this year that I finally [felt] like we were going in the right direction.”

As of the 2014-15 academic year, Belleworth had enrolled more than 500 students, almost all of whom were Latino, the remainder being African American. Over 25% of Belleworth’s students were classified as ELLs, and 95% of all students were considered socioeconomically disadvantaged by District measures. Ms. Heredia, the incoming principal for Belleworth, looked forward to the year as an opportunity to improve teaching, learning, and support programming for Belleworth’s student population, considering herself an advocate of using data to inform school strategy. Indeed, her understanding of how to use data to guide school-based decision-making was one of the key skill sets faculty considered in selecting her as principal. Ms. Nava, a teacher on Ms. Heredia’s hiring committee, summarized the faculty’s interest in retaining a data driven principal:

*So basically we just needed somebody... we needed someone to SHOW us, show us HOW you can use data to improve your school AND what KIND of data you should be using. You know, everybody can give a test and say, OK, based on the scores of these tests, we’re going to know what to do. It’s NOT about that.... So for me, it’s not just looking at that, it’s looking at... you have to look at grades, you have to look at, you know, everything. Everything needs to go into... those decisions.*

As depicted by Ms. Nava, Belleworth’s faculty were looking to new leadership for guidance and capacity-building in understanding how to use multiple forms of data in order to direct school improvement. Faculty knew using data to improve school performance was a key focus area for Belleworth, but they lacked the experience, practice, and technical knowledge to self-initiate these processes. Important to Ms. Nava, as a member of the hiring committee, was

finding a principal who understood and could build Belleworth's capacity in gathering and taking stock of a portfolio of data sources as a way of informing those improvements.

Belleworth, like the other cases comprising this study, has experienced its own share of growing pains in establishing itself as a pilot school. Now in its fourth year, its mission and vision are more solidified. With operational structures now more firmly in place, administration and staff are initiating a more systematic approach to the use of data in making decisions around academic programming. This can be seen in the analysis and reporting of student grades to faculty at regular intervals throughout each semester: time is set aside for grade review and interpretation as needed in weekly professional development meetings inclusive of all faculty, and student performance data are brought to the attention of the Governing School Council. These activities, however, rely upon the initiative and effort invested by motivated individuals (such as the principal, or teachers who have the technical skills to draft and analyze school grade reports), and are not yet entrenched into Belleworth's regular rhythm, schedule, and academic timetable.

The example of Belleworth, therefore, provides insight into how a school transitions into the regular, systematic use of data to inform teaching and learning, and what it looks like to establish such routines. In addition to intentionally selecting administrative leadership to guide data use, Belleworth was looking to develop faculty capacity in supporting and substantiating data-based activities. As they considered which of its objectives to measure, as well as how to measure them, administration and staff also needed to determine how to cultivate effective discussions around data analysis and interpretation. Additionally, as Belleworth responded to both internal and external demands for data and evidenced-based performance, this school

presented a case of how both classroom-based and school-based data are weighed, prioritized, and incorporated into a school's instructional decisions.

### **Case #3: Woodson College Preparatory School**

The establishment of Woodson College Preparatory School as one of the District's very first pilot schools was dependent on substantial public debate and District policy development over the course of several years. Additionally, with a deep commitment to ensuring students were well-prepared for college, Woodson involved the faculty and administration from a local institution of higher education as part of its design team. Once the District determined that the pilot school model would be endorsed, and following careful consideration and a feasibility study conducted by its university partner, the Woodson design team launched into their work of bringing the pilot school to life. Partnering with a university would not only ensure a "brand name" endorsement of Woodson within its local community, but would also provide Woodson with important resources, such as technical advisory and research personnel, from a top-tier research university. Eighteen months of proposal writing conducted by university personnel, the incoming principal for Woodson, and three lead teachers, were coupled with the guidance of a high-profile advisory board to produce Woodson's final structure, mission, and vision in 2009.

The notion of using data to empirically monitor and evidence Woodson's performance was infused into even the earliest conceptions of its design as a "best practice." Dr. Baher, one of the leading founders of Woodson, recounted the expectations expressed by the university in establishing the partnership:

*Well, early on there were VERY explicit conversations, both at the advisory board level and in other meetings, that the school first and foremost had to be a success for students.... That meant that... we had to have GOOD measures of learning progress. We had to have very good data on college going.*

From the outset, Woodson knew it would need to identify and track indicators of student academic progress as evidence of its own progress and development. An early emphasis on such data meant that systems and structures for data collection, analysis, and interpretation would need to be developed alongside all other operational procedures. School-based data collection could not take a backseat while Woodson gained its legs.

Determining what those early indicators of performance should be was no easy task, and while this was certainly an expectation of its university partner, Woodson's founders also expressed the need to uphold a sense of accountability to the District on its own. Dr. Baher recalled some deliberation over Woodson's participation in state exams, for example. One board advisor suggested that Woodson obtain a waiver from the standardized assessments, suggesting, as Dr. Baher recounted:

*This school will be at RISK if these test scores are the focus of its accountability as an early, fragile school. That the idea of MAKING that the focus, we know, with this particular community... that this will ECLIPSE good work and derail innovation (emphasis original)*<sup>4</sup>.

On the other hand, another board member – and prominent player in the establishment of the pilot school model – felt that to abstain from standardized testing would appear as if Woodson was taking advantage of its position as a university-endorsed school. Dr. Baher remembered this advisor countering with “No, that's not strategic. We can't be the ones that don't play the same game... as everyone else. We *need* to have those measures... just like everyone else does.”

---

<sup>4</sup> N.B., all emphasized text in quoted material is original unless noted otherwise.

Such voices emphasize how Woodson, as a model of school reform, needed to weigh very carefully the public measures to which it would commit itself as evidence of its health and performance. Standardized test scores, on the one hand, represented a high-stakes measure of student academic achievement to which every school in the District was beholden. Not to participate in state exams could have been interpreted as a signal of Woodson's belief it was an exception to the rules, unfairly benefitting from its university partnership. On the other hand, state test scores also served as a limited representation of school performance and were perceived as a potential threat to Woodson's work as a new school. Initially, low test scores resulting from early innovations might have underplayed Woodson's substantive progress and become, however unintentionally, the focus of negative attention.

Ultimately, Woodson decided not to apply for a state exam waiver. While it is unclear as to what benefits this returned to its reputation as a member of the pilot school community, Dr. Baher noted that test data unfortunately became "a huge source of strain, and anguish, and *stress*." In large part this was due to Woodson's rapid growth. Woodson's designers had intended for the school to start by offering only grades K-5, gradually including additional grades in subsequent years until it could offer placement to 600 students in K-12. In implementation, Woodson began as a K-5 school, started its second year as a K-11 school, and offered classes in K-12 its third year of operation to over 1,000 students. The ways in which Academic Performance Index (API) scores were calculated with the sudden addition of Woodson's "upper school" seemed to suggest that Woodson's state test performance dramatically decreased in this time of expansion.<sup>5</sup>

---

<sup>5</sup> Woodson suggests that this drop in its API score was an artifact of combining the school's new upper grades scores with the lower grades scores. Statewide, high school API scores were observed to be more than 70 points lower than

By state standards, Woodson was automatically classified as a “failing” school and, as Dr. Baher recalled that “there was this *massive* confusion and... you know, bad feelings and stress around that. And that was a *really* pivotal moment. We were a failing school all of a sudden, overnight.” Much work had to be done to convey to Woodson’s stakeholders (and, particularly, its university partner) how to understand and interpret its test scores in the context of its development. “Very early on,” Dr. Baher described, “it became *crystal clear* that that could not be our main narrative. We had to have *better assessments*.”

Woodson, led by Dr. Baher, began to engage in a process of developing its own student assessments which could contribute meaningfully to teachers’ in-class practices and serve as an additional metric of student performance. The school began by focusing on reading assessments in the “lower school.” This process of development, trial, and implementation involved substantial time and effort on the part of teachers and staff; but, in the end, as Dr. Baher summarized:

*This [was] our assessment. We’re going to the mat for this one. That was so important. That was the thing that, if I had to go back, I’d say, 100% do that again. Spend all that time and energy worrying about what assessments, because... that grounded people’s sense of ownership over the measures that would be used to gauge their progress.*

Building a systematized approach to student assessment for Woodson was guided by teachers, supported by resident “expert” staff, and intentionally leveraged Woodson’s assessment autonomy from the District. These factors were viewed as essential in cultivating a sense of

---

elementary scores, a difference that explained Woodson’s apparent second-year API decline (“Talking Points on Woodson College Prep’s 2013 API Score Drop,” 2013).

faculty value for the assessments, a key component in establishing their perceived validity and use. Subsequently, assessment development expanded into the upper school (a focus of Chapter 7).

While alternative assessments were an early target for Woodson, they had been complemented by a host of research and evaluation activities undertaken by various school stakeholders. Summer research conducted by faculty members, external studies conducted by partner university researchers, surveys implemented by parent groups, and research conducted by student groups were just some of the school-based data collection activities ongoing at Woodson. Weekly time for teachers to meet with colleagues and confer about student performance data was built into the school schedule. Release days made up a substantial line item in Woodson's annual budget so that teachers could hold intensive meetings around the development, implementation, and scoring of student assessments. Parent committees regularly reviewed school data with the principal, and a formal report of school progress, consisting of several labor-intensive measures of school performance, was annually produced for review by Woodson's university partner, as well as by the general public.

Woodson had clearly striven to develop a robust system of data use from even its earliest days in design. Certainly, this prolonged commitment to providing empirically-based evidence of its successes and challenges resulted in a very strong infrastructure of data use. This is not to say, however, that Woodson was immune to struggles in implementing fluid, school-wide data routines. As one founding teacher remembered:

*When we opened this school, there was absolutely NO data. And we had kids from 55 feeder schools.... We had NO information on them because [cumulative files] take forever to get sent to a school....*

*Up until the doors opened, we weren't exactly sure who was going to walk through the doors because of the... process of how students enroll. And so... I mean, we had SOME indication of who would be coming from [the District student information system], but... there was just so much work at the front load of opening the school of like, getting the programming done and what not that... I don't think teachers actually had an opportunity prior to the school opening to really look at... who's in my class, right?*

An advocate for the use of data to inform teaching and learning, this teacher had seen Woodson mature from its state of infancy – when data were either not available or not yet integrated into teachers' daily classroom activities – into a school that had increasingly routinized the use of student data into conventional practices. Part of this, she emphasized, had to do with amassing disparate data pieces (such as from District information systems, physical student files, and school-based data collection activities), and establishing Woodson's infrastructures to support their review, analysis, and interpretation.

Importantly, there is a distinction between designing and supporting a data use infrastructure and grounding it as a real foundation of school practice. This particular teacher attributed the adoption and evolution of data practices by Woodson's faculty to the earnest efforts it had made in building staff capacity to actively use available information systems, as well as to cultivate their own personal connections to the data. As she put it:

*So I think a lot of it at the initial outset [was]... It [had] a lot to do with the capacity that you need to build among the staff to use the information systems. Because the information systems always exist....*

*And it's not just like numbers, you know, it's things like... even some of the uncountable student comments, or parent comments, or things like that. So I think that there's just this... huge MOUNTAIN of data that exists, and so... I think in order to really USE it is... It really has to start with, like, the people who are there, and the things that matter to them.*

In the eyes of this founding teacher, Woodson's school data had "always" existed. Taking into consideration what data are housed in District information systems and what potentially less-formal data are collected from school-based activities, there is more than enough information for schools to draw on as evidence of its work and performance. While in its earliest days, data weren't always accessible to Woodson faculty, over time Woodson had strengthened its underlying infrastructure of gathering and collecting data, as well as processing, analyzing, and reporting school-based information. What embedded these systems into practice, this teacher argued, and what cultivated a genuine use of data, was the way in which Woodson data engaged faculty, whether they spoke to the "things that matter[ed] to them," or not. In this way, data use processes at Woodson was primarily teacher-led.

As of the 2014-15 academic year, Woodson College Prep enrolled just over 1,000 K-12 students. Of these, about 400 students comprised the upper school (Grades 9-12, the classes of focus for this research). Across the entire school, nearly 90% of Woodson's student population was considered "socioeconomically disadvantaged," and nearly 45% were categorized as ELLs. The majority of students (around 80%) were Latino, while another ~10% were Asian.

Woodson College Prep, with its strength in evaluation and research resources, continues to immerse itself in a dialogue about what it means, and what it would take for teachers, to use data "authentically" in its practice to promote student achievement. The case of Woodson thus presents a valuable view of how a relatively robust system of data use is embraced by faculty with various orientations to data, as well as the types of challenges teachers face in creating, managing, and analyzing their own measures and data use processes.

## **Cross-Case Insights**

Given the site-level flexibility afforded to pilot schools, it is important to understand the contexts in which their data use systems and structures are embedded. This is because the vision behind each school, and each school's mission-directed approach to school administration, teaching, and learning, are the impetuses for a school's orientation to, and value for, data. School infrastructure to gather, house, and disseminate data, as well as the policies, procedures, and routines guiding data use, are established in response to self-determined pilot school objectives. The foundation of data routines and infrastructure is imperative to the systematic use of school-based data in making decisions related to those objectives.

While systems and structures of data collection, analysis, and interpretation are important in guiding effective data use, they are not necessarily a prerequisite to data use. Data can be used to inform school-level decision-making without formal systems and structures in place, or while they are being developed and refined. This study focuses on three pilot high schools in various stages of growth and development, and explores how data use takes place with or without the institution of data routines and infrastructure. While The Academy is working through challenges of compiling data and has not yet routinized systems of data review, Belleworth School of Arts and Technology is wading through what it means to regularly review data and translate them into action. Processes of data use were integrated into Woodson College Preparatory School's original design and, years later, were still undergoing revision and refinement in response to stakeholder needs.

Each case within this study responded differentially to its unique school communities, inquiries of practice, and political-organizational spaces. This chapter is a first step in understanding the components underlying schools' orientation toward data use. The varied and

specific contexts embodying each school suggest that the three cases are not reasonably compared in terms of data use “proficiency.” Rather, the cases present a nuanced perspective on how data use takes place in three different scenarios. The next chapter continues this discussion examining the ways in which decision-making processes play out within schools, and the determination of who is responsible for making school decisions, plays a substantial role in whether, and to what extent, data are actually used.

## **CHAPTER 5**

# **CULTURES OF DECISION-MAKING**

### **Introduction**

Alongside the development of systems and structures to support effective data use in schools, the cultivation of decision-making processes and procedures is required to institute a demand for data among stakeholders. Data and data use processes are developed in this way to respond to questions of practice and organizational management as posed by decision-makers. There certainly is no shortage of decisions requiring the input and action of teachers and principals in their classrooms and schools. How these decisions are addressed, however, and by whom, are culturally-defined characteristics of a school. Subsequently, this chapter investigates how decision-making processes and relationships among decision-makers within each school case influence the degree to which school-based data are referred.

In accordance with their memoranda of understanding, pilot schools do follow general District guidelines on school governance. All cases within this study have thus established a Governing School Council comprised of elected teachers, the principal, parents, and students, and an Instructional Leadership Team composed of teachers, as their primary decision-making bodies. But not all decisions are formal, and a great deal of discretion is left to schools as to how they engage in dialogue, discussion, and day-to-day management of school and student issues. As a result, decision-making processes within schools are often found to be fluid, non-linear, and vulnerable to constraints in time and resources.

At The Academy, for example, the gradual institution of procedural decision-making was examined in the context of the school's first years of foundation. While concerted effort was invested in its establishment of a "flat" hierarchical structure, wherein teachers are given both

administrative authority and responsibilities, immediate organizational needs were often seen to take precedent over consensus-driven deliberation. Data, in this case, were valued by decision-makers but did not yet factor into decision-making processes still under construction. Belleworth School of Arts and Technology's new principal was hired for her strength in using data to inform school strategy. However, the use of data to drive instructional change was also accompanied by changes in the allocation of decision-making authority. The case of Belleworth provides important insight into how data are an element of school politics rather than a standalone resource upon which decision-making draws. Finally, although Woodson College Preparatory was regarded to be very strong in its approach to data use by design, the early experiences of its principal shed light on the importance of defining leadership roles and brokering personal relationships with individual faculty before being able to engage in discussions involving school systems of data use. In view of these experiences, while this study initially sought to investigate how different types of data influence the variety of decisions made within each pilot school, it was found that understanding the establishment of decision-making processes was a prerequisite to understanding their propensity toward data use.

### **The Academy: Real-Time Decisions and Aspirations of Data Use**

The Academy's proposal to become a pilot school documented its intended curriculum, pedagogical approach, governance structure, and overarching policies and procedures. Once the school opened, however, The Academy's design faced the challenge of implementation. For The Academy's first time administrators, which included all teachers in its "flat" managerial structure, this entailed a steep learning curve in terms of both operational and instructional planning. Indeed, "planning" was not nearly the orderly act of faculty sitting down and mapping

out strategic objectives, their related activities, and the infrastructure and processes needed to support those activities. Rather, planning seemed more to be the substance of professional development sessions scheduled twice weekly, organized meetings with the formally-elected Governing School Council and Instructional Leadership Team, and impromptu discussions between founding members of the school.

### **Decision-Making: Form vs. Function**

The Academy took seriously its focus on consensus-based decision-making among all faculty members, as well as the responsibility of all teachers, in managing the school's administration. But while these principles continued to imbed themselves into The Academy's day-to-day functioning, the shortage of time and staff availability to plot and plan remained a constant obstacle. There existed a pervasive feeling among faculty that informal decision-making was undertaken by a consistent core of faculty as a default mechanism, and this perception had been a threat to promoting a collective sense of governance. Mr. Knowles, a teacher from The Academy's design team, explained:

*Because, you know, the problem is that... Rob, Mr. Cooper, and I were some of the earliest people that were working on this. And, we get together and talk about stuff, and people think like decisions are being made. Like one day Rob, Mr. Cooper, and I are talking about international politics.... But still there's this perception that things are going on behind closed doors, and that SECRET things are happening and "they're not being transparent, and they've just decided this thing without us." And there's all these processes and systems in place, and there's still this, "us versus them" mentality.*

Mr. Knowles noted an atmosphere of distrust developing around the frequent conference of this small group of The Academy's founding teachers and principal. He sensed that his casual conversations with colleagues were interpreted as "secret" meetings wherein critical decisions

were made without transparency. While these former members of The Academy’s design team might once have functioned as the school’s primary leaders , Mr. Knowles suggested that this authority had devolved into proper “processes and systems” of whole-school decision-making. Mr. Knowles attempted to reinforce this message by reminding his peers during professional development meetings that “everybody here is an administrator. Think like an administrator, because you’re doing administrative-like things.” He believed that this consistently reiterated message was slowly weaving into the faculty’s fabric, as confirmed by his colleagues’ increasingly vocal recapitulation of the idea.

The principal, Mr. Cooper, also emphasized The Academy’s attempt to do away with the conventional dichotomy of “administration vs. faculty,” or a culture of “us versus them.” He made continuous attempts to remind teachers that “the four guys in the Governing School Council” – three of whom had been affectionately dubbed “the triumvirate” by The Academy’s remaining faculty – did not constitute a pyramid structure of top-down management and was not The Academy’s model. That all teachers had an equal stake in the school’s administration and governance was one of the central pillars of The Academy’s philosophy.

Another mechanism employed by faculty to bolster an understanding of cooperative governance was what The Academy termed “The Fist of Five.” This was used as a standard meeting protocol to facilitate discussion and “take the temperature of the room” in the process of decision-making. As he held up one hand to demonstrate, Mr. Knowles explained:

*And the fist is, I absolutely wouldn’t, it violates my core principles. And one is, I’m not going to do it, you have to convince me to move off of one. Five is I will... and this is kind of... the clarifying point is, 4 means you’re completely for it and you’ll do everything you can. Five is you’ll RUN it. And so I want to clarify, because people will hold up fives and I said, you mean you all want to run it!*

Using the “Fist of Five” convention, consensus among faculty was sought for all decisions brought to the group. The Academy prided itself on its work to integrate all faculty voices into its final decisions rather than to rely on majority vote, which the principal believed could result in demoralized subsections of outvoted faculty.

Despite its focus on a “flat” rather than hierarchical management structure, and the integration of all faculty in school wide decision-making, it was clear that the pace and immediate operational needs of the school outstripped The Academy’s capacity to engage its entire faculty in every decision. This applied not only to day-to-day issues of school management, but also to fairly substantial decisions concerning budget and programming.

Take, for example, The Academy’s management of Title I funds. Due to The Academy’s unexpected faculty cuts last year, Mr. Knowles, who was originally designated as a part-time Title I Coordinator, found himself teaching a full load of classes with little residual time to learn how to navigate the roles and responsibilities of the coordinator position. Fortunately, the District’s practice of assigning temporarily displaced teachers to attend work at schools (a/k/a “pool teachers”) allocated an aspiring administrator to The Academy. This temporary faculty member volunteered to spend his time coordinating The Academy’s Title I funding. As a result of his work, The Academy discovered that they were eligible for nearly one-third more funding due to its proportion of enrolled low-income students. This resulted in a budget windfall in the middle of the academic year. Mr. Knowles recounted:

*Middle of the semester, this money, poof, magically appears.... Which is a HORRIBLE way to do things, because then you have to spend the money without a plan, basically. You have to INVENT a plan to spend the money because it’s going to go away!*

With no plan in place, and money needing to be spent immediately at the risk of losing it entirely, Mr. Knowles, also chairman of the Governing School Council, called an emergency meeting. Quickly, during lunch, a gathering of enough Governing School Council members needed to make quorum was assembled. The group wrestled with how to propose appropriating the funds in time for a vote at the next scheduled Council meeting at which faculty, administration, parent, and student members could weigh in. However, given the complexity of the budget revision, combined with an urgent timeline, Mr. Knowles eventually suggested granting the Instructional Leadership Team the authority to act on behalf of the Council. This team was comprised of Governing School Council faculty members and the principal, but did not require a formal vote from its members, nor the input of parents and students.

Mr. Knowles readily admitted that the form and function of these budget allocations were not ideal and that the Council was “called to task” on its decision-making process:

*Since I’m chairman I’m supposed to develop the agenda and distribute the information.... The budget has been developed by Ana and myself with input... that was the first year. And then it was Ana and Mr. Cooper, and whenever people [had] the chance to talk... that’s how it’s developed. And it’s basically brought to the Governing School Council already done for approval.... And it was supposed to be more like, developed by the Governing School Council and then given to the SCHOOL rather than by the school and then given to the council.*

*And so... because we’re just treading water and we’re trying to do all these things, and there just hasn’t been time to make it work the RIGHT way. OK this will WORK. But one of the parent members [said], “I’d really like to SEE and help out... and understand this.” And so that’s one of things that we need to work on this COMING year, is... bring IN the parental and student voices.*

From the identification of funds central to the full operation of the school to the decision-making processes designating those funds, this example illustrates The Academy’s difficulty in simultaneously founding decision-making structures while receiving incoming flows of new

information. In this case, the urgency of needed budget appropriations superseded a process of input from all stakeholders which would have presumably entailed a broader discussion of school objectives and priority activities for Title I students. At stake, however, was a substantial portion of funding essential to the benefit of The Academy's lowest income students.

As the chairman of the Governing School Council accentuated, the pressing demand of functionally getting a host of things "to work" had sometimes trumped The Academy's ability to "make [them] work the right way." In addition to issues of budget, for example, was the creation of teacher committees to oversee other administrative functions such as personnel management. Although a teacher committee was charged with overseeing the recruitment of new teaching personnel, they were not actually involved in The Academy's substantial hiring wave at the end of the academic year. Mr. Knowles, who was coordinating the hires, hadn't realized the teacher committee was actually meeting, admitting that a lack of communication regarding the committee's work that year precluded its engagement. An example of independently-driven decision-making might be Mr. Cooper's self-initiated arrangement of a spring semester "reorientation" for all students following their return from Winter Break. On the students' first day back, Mr. Cooper arranged for each grade to split into multiple groups and attend a rotation of teacher-led sessions revisiting school policies and procedures, students' personal academic goals and strategies, and graduation requirements. While Mr. Cooper considered the day a success, he admitted that "the staff was completely blindsided by this" and that they were given less than an hour to prepare the "reorientation" before it went into implementation. Mr. Cooper's intentions were to quickly mobilize an innovative approach to the second semester and to avoid forcing his staff into extra work on their break. While none of the teachers interviewed expressed

resentment at their late involvement in the event, one teacher mentioned that the apparent lack of planning likely impeded the desired impact of the day on students.

### **Disparate Data Use Activities**

Although The Academy may still be actively establishing decision-making systems and procedures, this is not to say that The Academy has not attempted to collect and compile school-based data to assist in substantiating its decisions. For example, Mr. Knowles – also considered The Academy’s “tech guy” – had been working all year to master the District’s online student information system. His goal of analyzing “D/F rates” as a metric of student achievement, however, had been compromised by technical difficulties. For whatever reason, at the time of study, the new system was unable to calculate rates of student failure. Mr. Knowles pointed to an old desktop computer he had erected in his classroom to access the District’s former data system as an alternative solution. However outdated, this technology allowed him access to students’ original grade records, from which he manually calculated D/F rates. While Mr. Knowles described this process as “not *that* hard,” he noted that it took “time to have to go through and make it all work.” Subsequently, his ability to tie the success of student resources and interventions (such as the introduction of laptop computers with Title I funds), to improved student proficiency rates was, at minimum, laborious. Additionally, he felt that for all of the effort he invested into producing the failure rates, they really only attended to The Academy’s Title I report requirements. To apply these data to other school purposes, Mr. Knowles suggested, was a “problem” for a new school with only a limited amount of time to thoroughly analyze the data.

Indeed, Mr. Cooper readily identified several other sources of school-based data as evidence of school performance, particularly in view of self-study materials required for The Academy's upcoming accreditation review. For example, he pointed to the administration of several student surveys as a measure of school climate. Their analyses were the responsibility of Mr. Easton who reportedly involved The Academy's student council in reviewing results. However, for personal reasons, Mr. Easton was not available for activities outside of teaching at the end of the academic year, and the extent to which those data were analyzed and then disseminated to faculty members is uncertain. Efforts to make use of these data, then, seemed to falter as Mr. Easton was unable to fulfill this auxiliary role. At the time of study, teachers also facilitated an extensive student survey to all senior students regarding their experience at The Academy. Two teachers were charged with entering the data into an online platform and producing an analysis for faculty. Technical difficulties with the online survey service, such as obtaining a paid license and compiling a single data set from several survey "parts," as well as ensuring all paper forms were entered for all parts of the survey, set analysis activities back by several months. It was unclear whether teachers had had the time to actually analyze the data and present the resulting findings. In both of these examples, how data were meant to contribute to a larger discussion of school performance is unknown. At the moment, the absence of perceived consequences or feelings of urgency associated with outstanding data analysis activities seemed to reflect The Academy's overall "need" for data. While valued, data were not necessarily a priority.

The Academy clearly intended to collect, compile, and somehow make use of data. However, it had not yet articulated an overarching plan as to how these data were expected to answer its various decision-making needs. The use of data to inform The Academy's progress

and development was not likely be seriously considered until it was determined who would be charged with data collection and analysis, for what purposes, and how and when those data would be reviewed and interpreted by faculty. This would have simultaneously entailed the systematization of decision-making processes. Faculty participation in collecting, analyzing, and using data to inform decisions made regarding The Academy's teaching and learning activities would have relied on their sense of worth in making those decisions. Until The Academy iron out the procedures by which decisions are made, it will continue to wrestle with how its faculty think and move as a collective. As such, systematizing basic protocol through which administrative and instructional issues may be reliably addressed seems inextricably linked to the integration of school data informing those processes.

### **Belleworth School of Arts and Technology: Power, Authority, and Then, Data Looking for Leadership in Data Use**

The question of who makes decisions at Belleworth and the ways in which those decisions are made became a substantial point of contention. On the one hand, Ms. Heredia's first year as principal was looked to as an opportunity for Belleworth to, in Ms. Nava's words, "strengthen its school culture" and to "set clear expectations for teachers and for students." She explained how, following the retirement of its founding principal, Belleworth was looking for out-of-classroom support with student discipline, instruction, achievement, and maintenance of its high expectations. As part of this, Belleworth's faculty were interested in finding a new principal who exhibited strength in analyzing and interpreting school-based data. Members of the hiring committee even built an exemplar activity into the interview process, as Ms. Nava explained:

*So now that we have leadership that understands the importance of data, and in fact, our faculty did, because one of our questions we asked when we were hiring our new principal was all about data..... We gave them the data of our school and one of our questions was, based on the data of our school, what can you tell about our school? What do we need to do? Because we knew, like WE as a faculty know, we NEED to look at data in order to figure out our school needs. And so that was one of our BIG key things of hiring our new principal was the fact that she understood.... And she was able to look at our data and tell us exactly what we needed to start doing.*

When asked what type of data had been given to candidates for review, Ms. Nava explained:

*(Sighs). Our EL population. The reclassification was very low. Our parent involvement, I mean we had like 1% parent involvement that took our compact survey. Our [exit exam] scores, even though our [exit exam] scores she said were very good, she was able to zoom right into that and show how the year before, she was like, "Oh it really increased, so whatever you guys did, we need to keep doing, and do it even better." So that was really good, you know, that she was able to do that.*

Ms. Nava emphasizes here that the faculty recognized a pressing desire to use school data not only to provide more effective support for students, but to also target school and student needs. Interestingly, Ms. Nava already seemed to have had some personal insight into the data, pointing out what major points Ms. Heredia was correctly able to identify in her interview. As such, the activity was designed to test Ms. Heredia's capacity and skill in data identification and analysis, as well as the alignment of her interpretation with that of interviewing faculty. Ms. Heredia's ability to pinpoint evidence of programmatic potential simultaneously exhibited her philosophical and leadership approaches to intervention and action and the ways in which these were guided by her read of the data. Ms. Nava continued her discussion of Ms. Heredia's selection:

*So basically we just needed somebody... we needed someone to SHOW us, show us HOW you can use data to improve your school. AND, what KIND of data you should be using. You know, everybody can give a test and say, OK, based on the scores of these tests, we're going to know what to do. It's NOT about that. You know, like, I'm not a test taker, I don't like tests. I don't like test taking because I was not a good test taker in my life. I was the one who gets stressed and... So for me, it's not just looking at that, it's looking at, you know, you have to look at grades, you have to look at, you know, everything. Everything needs to go into those decisions.*

As depicted by Ms. Nava, Belleworth's faculty were looking to new leadership for guidance and capacity-building in understanding how, exactly, to use data to direct school improvement. In other words, while faculty knew this was a key focus in upcoming years, they lacked the experience, practice, and technical knowledge to self-initiate interventions in response to the data. Important to Ms. Nava was finding a principal who understood the significance of gathering and taking stock of a portfolio of data sources. This was fueled by Ms. Nava's own personal disregard for high-stakes measures of school performance, such as standardized testing, which she found both insufficient in portraying whole-school quality and not representative of Belleworth's vision. A principal focusing on performance indicators lacking faculty endorsement was not anticipated to be a good fit for Belleworth.

### **Learning How to Leverage Data**

Ultimately, Ms. Heredia's introduction into Belleworth has proven, for several teachers, to be a great benefit by way of engaging the school in a "new culture." In a separate interview, Ms. Nava went on to explain how Ms. Heredia's modeling of data-based practices had re-oriented Belleworth's approach to student support:

*So we just got Ms. Heredia, and she's amazing so... the whole culture of the school has changed a lot. So before... we talked about instruction but it was just*

*kind of like, “Oh my gosh, these kids aren’t doing as well as we want them to.” But it wouldn’t move forward from there. And so now it’s like, “OK, this is how many of our students are doing bad,” and I think that’s part of [our Response to Intervention Committee] and part of her leadership.*

*....So that’s how our [professional development sessions] were run before. It was kind of like, a meeting of the minds and... it would be like a discussion and then it wouldn’t... move too much further from where the discussion was. Whereas now, I think with [Response to Intervention Committee], and even the [professional development] that we’re doing... It’s like, we see that our kids are struggling with reading. So OK, so now what are we going to do to mitigate that? So we brought in different programs, different people. It’s been good.*

The kind of progress Belleworth has made in using school-based data for instructional improvement was refreshing for Ms. Nava. Whereas previously faculty would discuss student performance at professional development meetings, their ability to move from dialogue to actionable next steps was a significant challenge. Diagnosing student need, or underperformance, was an important conversation, but a seemingly intractable challenge. While Ms. Nava did not point out which specific strategies Ms. Heredia had used to propel the faculty forward from the diagnostic stage (apart from infusing data-based discussions into professional development meetings), Ms. Heredia’s self-depiction of her own leadership style suggests little hesitation to initiate intervention programs and faculty committees or to endorse larger scale structural changes such as the implementation of an entirely new bell schedule. Ms. Nava also makes reference to Ms. Heredia’s help in launching the Response to Intervention (RTI) initiative at Belleworth, which has become a systematic approach through which Belleworth’s faculty collectively review and interpret student performance data as a way of informing curriculum and instruction (see Chapter 7). As a result, the diagnoses of specific student need areas (e.g., improved reading performance) using data are now integrated with discussions of how faculty

can appropriately respond to these needs and by the consideration of follow-up resource provision (i.e., “different programs, different people”).

Ms. Heredia’s ability to effectively mobilize Belleworth’s faculty in response to specific student and school needs stands alongside her long-term strategy for Belleworth’s growth and maturation. She explained that her own vision for the School and the ways it uses data to guide instruction and programming necessarily entails time for meaningful implementation, observation, and reflection:

*So I think right now is just like, putting down...the foundational pieces, you know? In getting people to have a common understanding of things. Like, what should be what we're doing? I think next year is going to be more of like, okay, let's start testing this out and taking it for a ride, and in the third year, it's probably going to be like, what can we do better? What could we improve? That's why we say... our five-year plan, because I think it's going to take a while.*

Like the principals from both The Academy and Woodson College Prep, Ms. Heredia recognized that establishing, revising, and refining school culture takes time. The School’s approach to instruction and service was a cultural orientation. As Ms. Heredia described, first there was the need to develop a “common understanding of things,” such that the faculty could collectively determine how Belleworth should prioritize its strategies and activities. Once a plan of action had been collaboratively established, she intended to dedicate the following school year to trialing and testing these approaches, and a third year to reflecting on Belleworth’s progress and plans for further improvement. With iterative cycles of revision and improvement, Ms. Heredia plotted the strategies implemented the past academic year on a five-year timeline. This, she asserted, was a reasonable period within which to evidence real school improvement.

Built within this five-year plan was a reorientation to teaching and learning practices within Belleworth. Ms. Heredia, for example, expressed frustration with the limited ability of her

faculty to be reflective on their performance and instruction. With a relatively new school and several teaching staff also new to the profession (4 of the 5 teachers interviewed from Belleworth have had five years or less of teaching experience), the frame of reference for Belleworth's faculty in terms of "excellent teaching," Ms. Heredia argues, was rather narrow. This insularity contributed to Belleworth's difficulty in being self-critical and identifying areas for instructional improvement. As a result, the quality of teaching she observed in Belleworth's classrooms the past year was not thoroughly exemplary. Ms. Heredia recounted how, upon her entry, faculty struggled to identify what Belleworth was doing well in terms of its instruction. When asked, faculty applauded its supportive staff relationships ("We love our friendships") and faculty satisfaction ("We love our low turnover rate with staff"), both being aspects of strong school cohesion and environment, but not of classroom instruction. Part of Belleworth's growth as a self-reflecting school, Ms. Heredia suggested, would be teachers' ability to have honest conversations with each other about their performance. While teachers will often be critical of each other in private, she observed, they neglect to raise these points in face-to-face discussions with their peers. Ms. Heredia was, therefore, forced into assuming the uncomfortable role of instigator and pushing her teachers into publicly raising their critiques.

### **Devolution, Dissolution, and Discord**

Ms. Heredia's own active leadership style, however, has not been met without some resistance. Even the establishment of a five-year strategic plan required substantial team-building and negotiation. In particular, Ms. Heredia's close watch over teachers who she felt were not serving the school's mission, as well as her perceived lack of clarity in communicating her school vision, created a sense of frustration and distrust among some faculty. This became such

an issue among Belleworth's teachers that they organized meetings to discuss Ms. Heredia's leadership and voted to send their Union Representative to engage Ms. Heredia in a formal conversation voicing faculty concerns. Having been forewarned of this by one faculty member, Ms. Heredia invited her Instructional Leadership Team to her house on a weekend, and over mimosas and collective dialogue about what the entire team would like to see accomplished at Belleworth, they were able to collaboratively map their prioritized outcomes backwards into the five-year strategic plan.

Following this team-building exercise, Ms. Heredia was under the impression she and the faculty were again on the same page. However, by the close of the academic year, it had become apparent that her entry into Belleworth also resulted in a shift in power dynamics not welcome by all teachers. Surprisingly, this was most true for her ILT. The ILT, as it turned out, had been largely tasked with managing Belleworth since its opening, and the previous principal habitually deferred all issues of instruction and management to the team. The ILT regarded itself as a standing committee (not subject to election), overseeing the Governing School Council, Belleworth's instruction, and all faculty sub-committees. Ms. Heredia, however, reorganized this structure such that the ILT became one of several committees and the Governing School Council became the primary venue of school-wide decision-making. This devolvement of power, Ms. Heredia believed, became a major cause of strain.

Disgruntlement resulting from the dissolution of the ILT as the epicenter of Belleworth's decision-making structure was compounded by Ms. Heredia's personal value for principal autonomy. For example, she felt that there were some decisions that should be left to the discretion and authority of the principal based on the specific needs, timelines, and demands of

school administration. She recounted a conversation with one of her ILT members about her leadership perspective:

*Can you guys come and ask me why I made that decision? Yes. Will I explain to you guys why I did it? Yes. Am I always gonna' go to you guys for everything? No. And I feel like THEY, as a pilot school – pilot school leaders – thought that everything goes through them. They had structured [it] where EVERYTHING goes through them. Budgets? No.... The leadership part of the teachers LEADING is that they're leading in different committees and not that they DECIDE everything that happens at the school. And I think that's where it's gotten murky, where they've taken it as like, "WE have to decide, WE FOUR have to decide." Who said? Who elected you four?*

A wide range of issues, spanning from the revision of the school logo to determining whether Belleworth should move toward the full inclusion of its Special Education students to defining the scope of work for a new assistant principal, all became points of contention between Ms. Heredia and the ILT. In some cases, Ms. Heredia found herself making executive moves without direct consultation from her teachers, at times because decisions required an extremely quick turnaround, and at times because she felt she was acting in the best interest of Belleworth's students. She felt that relying on the ILT to make all decisions at Belleworth lacked a broader sense of accountability, particularly if the ILT insisted on representing the voice of all faculty without being subject to faculty vote. Ms. Heredia recognized that, while members of the ILT team had been pressured into positions of leadership under the previous administration, their understanding of all the "influences that shape decision-making" remained limited. In particular, Ms. Heredia had been frustrated by the ILT's propensity to make decisions she characterized as self-serving sometimes. For example, ILT decisions around the master schedule first took into account the preferred schedules of ILT members. Another ILT member refused to allow a new assistant principal to observe her classes and wanted to embed this refusal into the position's

scope of work. Ms. Heredia felt other members had derailed the interview process for the assistant principal because of their own personal issues with some of the candidates.

In contrast, while her own decisions may have been interpreted as draconian at times, Ms. Heredia pointed out:

*I mean, I sit back and I tell [members of the ILT]... “We laid out all the things we’ve done this year.... What has been bad for the school? What has led to negative outcomes for the school?” I don’t think they can find any, really, that have led to impacting students in a negative way.... Like all the things that we’re deciding on doing are things to improve and make the school better.*

*That’s why I feel like it becomes just adult issues. Because none of it is negatively impacting the students. Well, I don’t want to say “none of it,” but you know, most of it is about up here (angling hand at eye height). The decisions we’re making and the... hurt feelings and all of that stuff.... But what about the work? That’s the priority.*

For Ms. Heredia, leadership at Belleworth meant undergoing heavy criticism from some of her faculty members. But, at the end of the day, she saw her strategies as both necessary for Belleworth’s improvement and in the best interest of students. She described herself as selective about the issues for which she fought and for which she could be flexible. Unfortunately, she felt that the resistance she was encountering from some of her teachers stemmed more from a “power struggle” rooted in “adult issues” rather than discussion of student need. The latter, for her, was “the work” that should substantiate their internal debates. Although teacher leaders at Belleworth were primed and ready to direct instructional and programmatic decisions with school-based data, as it turned out, who makes those decisions – rather than what evidence is used to make those decisions – became the issue central to decision-making that year.

## **Woodson College Preparatory: Causal Relationships Rooted in Personal Relationships**

### **Internal and External Perceptions of Data Use**

Like The Academy and Belleworth, paramount to Woodson College Prep's experience in data use was the role teachers played in school decision-making processes. Woodson prided itself on the strength of being a "teacher-powered" organization. As part of this paradigm, lead teachers (one per department) were paid an additional stipend to be the instructional leads of Woodson, and one component of their job description was to facilitate professional learning. Analogous to "heads" of their department, the Lead Teachers were an essential terminal of communication and feedback at Woodson, maintaining the "pulse of the school," as it is described by one staff member.

When Ms. Figueroa, Woodson's current principal, first arrived two years ago, it took some intensive negotiation before she was able to leverage the strength of the school's teacher network in using data to guide instructional changes. Reflecting on her own entry into Woodson as its second principal, Ms. Figueroa remembered how her own vision of data use conflicted with that of the faculty.

*I [felt] like we didn't do enough... work around looking at student achievement. Like, actual achievement. And then ACTING on what we knew about that achievement in a systematic way. And I was surprised, because based on what I had read... the narrative in the annual reports or whatever, I felt like....*

*What I had done in other schools, which was set up these professional learning communities and I had, you know, we had SPENT money and resources on professional development, like, HOW DO we look at data? HOW DO WE improve our practice?*

*That even though that was there in THEORY, I didn't necessarily see it in practice. What I felt I found when I asked questions was about, OK, so what are*

*we doing with that? So we know that 50% of our students are NOT reading at grade level, what does our response LOOK like? I feel like no one could really articulate that clearly to me. But people were... uncomfortable with those questions initially....*

*Their responses were always very defensive about [it]: “But, but you know, we’re doing these other things” or “It’s because we don’t think of intervention in that way.” They were giving me their own definitions of things, but they couldn’t really articulate what that intervention, or what additional support LOOKED like.*

*And then I couldn’t find clear evidence that SOME ONE or a body in the school was really monitoring achievement. You know, other than like Dr. Baher’s annual reports, really INTERNALLY. Like people owning their own data and saying, oh yeah, this is ours and here are the ROOT CAUSES for that. Because at my other schools, those were the words we used: “What are the root causes of this under achievement that we have control over?”*

Coming into her position as principal, Ms. Figueroa drew an impression of Woodson’s strength in data use from the annual school reports produced by Dr. Baher. These rather comprehensive reports were publicly available and highlighted the school’s progress and achievement using several sources of student, teacher, and community data. Ms. Figueroa also brought with her experience using data for school-based decision-making from her previous principalship which had specifically targeted resources and teacher training toward the identification, analysis, and interpretation of data in the context of instructional practice.

Ms. Figueroa’s expectation was that Woodson’s teachers would be equipped with a framework to tackle questions concerning student achievement, such as understanding why 50% of their student population was not reading at grade level. In contrast, what she encountered was teachers’ reticence to “own” student achievement data in ways that showed a sense of internal accountability to the results. Rather than considering the “root causes” contributing to signs of student “underachievement,” Ms. Figueroa felt the teachers at Woodson drew up a defense in

how the data were interpreted. This was exemplified by some teachers' propensity to provide alternative definitions of what "intervention" meant in implementation and an ultimate inability to articulate what factors they accepted control over in improving student achievement, i.e., what additional support might actually "look like" in concrete terms.

Over and above a hesitation to look introspectively at their students' academic achievement data, Ms. Figueroa recounted the adverse emotional reaction Woodson's lead teachers had to this type of data use:

*I remember one of my meetings with my Lead Teachers... pretty beginning, maybe the second or third month [after I arrived]. I introduced... made some copies of tools I had used at other schools that had a data-driven kind of cycle. Like, OK, how are we incorporating a REAL cycle of... OK this is what we know about students, what piece of this are we going to try and address and improve? How did we know if anything happened? And if it did, if it didn't, what's our response?*

*And people were actually very... they just kind of like, shut down, and were like, you're not going to change what we're doing already. [Clicks teeth] Like, why are you bringing this? Are you saying now we HAVE to do this?*

Here, the introduction of "data-driven cycles" by Ms. Figueroa was viewed by the lead teachers as something extraneous to the current efforts of the faculty. In Ms. Figueroa's eyes, the purpose of using student data to identify strategies for improvement and the tracking of student progress was neither understood nor remotely desired. Rather, she saw the lead teachers "shut down" in light of her suggestions, offering in return only immediate pushback. Ms. Figueroa's suggestion to implement a new data routine was perceived as an external mandate insensitive to the decision-making systems and processes already established at Woodson.

## Building Rapport and Gaining Perspective

In her reflection on these early moments, Ms. Figueroa saw several factors contributing to this initial exchange. Part of this was Ms. Figueroa's own limited understanding of what emotional toll recent organizational shifts had taken on Woodson at the time of her entry. The founding principal had left for a job at the District, an interim principal had filled her spot only as a temporary place marker, and, due to budget restrictions, half of the staff had been let go and about one-third hired back by the beginning of the school year. Ms. Figueroa arrived in October, already three months into the academic year, and a sense of community and trust needed to be re-established at Woodson. In only its fourth year of operation (having quickly expanded from a K-5 elementary in Year 1 to a K-11 in Year 2 and a K-12 in Year 3), this was a challenge.

It was suggested to Ms. Figueroa that she first start with the lead teachers. She explained:

*I decided, OK let me step back, you know. Some people were telling me, get to know the lead teachers better, you know, get to build more relationships with them. Here at this school everything goes through them. And people respect them a LOT and they look up to them. And then remember that they're transitioning into new leadership. So the only people [the faculty] trust are the lead teachers who have been with them.*

Taking on this advice, Ms. Figueroa's next strategy was to suggest the observation of teachers' classrooms as a way of understanding how goals set at the department level through their self-defined "professional learning plans" aligned with instruction – an important link in understanding how indicators of classroom performance were construed. "Oh my God," Ms. Figueroa recounted, "That unleashed the second wave of like, NO! BECAUSE, in people's minds, this was tied to evaluation." Although Ms. Figueroa viewed classroom observation as an opportunity to see teachers in action and to become a partner in instruction, she discovered that the lead teachers perceived observation as an encroachment on their autonomy and, again, an

imposition of external accountability. Trust was not something that could be forged by a closer understanding of teacher goals and instructional execution. Ms. Figueroa found that this would only come by making personal connections with the staff. She continued:

*So people were like, well maybe you should have lunch with all of us. So I did. I had a couple of lunches with the lower school and then I started doing these little individual check-ins with people just to get to know them and see what they wanted and felt. By then it's like December, January. And so I started getting, like, a better sense of who's who. And again, just feeling like, OK, so I'm getting to know....*

*I start[ed] sitting in on their release days as departments and grade levels, and it was great because I was learning what they were doing. And I was very much in awe of the work they WERE doing. I felt like, OK, this is GREAT. I mean, because I got to sit in on ALL of them, and I started getting a better picture of the whole school.*

Through informal lunches, “check-ins” with individual teachers, and sitting in on teacher meetings held on release days, Ms. Figueroa was slowly able to gain a sense of teachers’ day-to-day work. By embedding herself in the mechanics of the school, and reaching out to each teacher on an individual level, Ms. Figueroa was finally able to acquire a better picture of the entire school over subsequent months, as well as much-needed insight into the work accomplished by Woodson’s teachers.

On this pathway, Ms. Figueroa made a concerted decision to table the issue of reviewing student achievement data and thinking about their implications on instructional strategy. There was a clear need for her to first make sense of the context of instruction underway at Woodson. Nevertheless, Ms. Figueroa was still agitated by the focus of faculty discussion on “This is how we *do* it, and this is how we *process*. But it wasn’t as focused on *outcome* and what causes that outcome, like in that causal analysis of, *why* do we continue to see this?”

## Securing Allies and Finding Pressure Points

Fortunately, Ms. Figueroa also began to forge a strong partnership with Dr. Baher who had been working with Woodson's teachers since the school's inception. Finding themselves like-minded in terms of how data could be used at Woodson College Prep, Ms. Figueroa identified in Dr. Baher an ally to help the school to make a transition toward "valuing" data, i.e., not just regularly collecting their own data, but also doing something purposeful with it.

The strength of Ms. Figueroa's relationship with Dr. Baher, who was herself well-known and respected by Woodson's teacher community, as well as Ms. Figueroa's own efforts to engage teachers on an individual level, were essential factors in her ability to gain credibility as a data use advocate at Woodson College Prep. Founding a sense of mutual trust with Woodson's faculty not only legitimized her position as principal, but was also imperative in resolving the otherwise fractious issue of engaging in conversations around student data. Previous research conducted on the introduction and evolution of student assessments at Woodson College Prep confirmed these findings (Quartz et al., 2014). They suggest that the constitution of trust among faculty and administration, as well as carefully-timed dialogue sensitive to the conditions of Woodson's teacher-powered professional development and instruction, laid the necessary groundwork to galvanize support for the school-wide assessment of student learning. Only after these pieces were in place was Ms. Figueroa able to assert a common understanding around the notion that "it's about really what we're trying to build together," and move the conversation of data use forward. She explained that she was then able to rely on the lead teachers to influence a cultural shift in data use:

*Because we create[d] PRESSURE through the leadership team, that mean[t] I'm a lead, I'm going to go back to my department. I can't just go back and then do*

*whatever.... The leadership team is pushing for something, we all do it together, and we go to our spaces and we PUSH[ED] for that.*

Through cooperative agreements as a leadership team, Ms. Figueroa believed Woodson's teachers were finally able to engage in discussions about activities like student assessments and classroom-based data use cycles. She recognized that there were still challenges associated with instituting a cultural shift toward data use at the individual teacher level. However, Ms. Figueroa had faith in the vision of Woodson's leadership team and its ability to "push," in contextually perceptive ways, for collective betterment. This was the foundation through which Woodson established a "teacher-powered" sense of buy-in, ownership, and mutual accountability.

### **Cross-Case Insights**

In its first years of operation, one of the primary objectives for The Academy was to build a "flat" management structure such that there existed no perceived segregation of power between teachers and administrators. Decisions were therefore based on faculty consensus, and The Academy adopted the "Fist of Five" approach to facilitate constructive negotiation and a system of decision-making based on dialogue. The school was admittedly still struggling with the production and review of standardized student and school data to monitor performance, but took pride in its focus on founding healthy relationships and a sense of school ownership among faculty and staff through its managerial approach. At the same time, because The Academy was a new school, it often felt pressured to make high-stakes decisions quickly. While it was perceived that data might be useful in these circumstances to help guide administrative decisions, The Academy frequently found that neither time nor personnel capacity was in adequate supply to sufficiently review and process such data. As a result, decisions could be made somewhat haphazardly, sometimes undermining its work to promote a culture of equal engagement among

all school stakeholders. Effective data use within The Academy was dependent not only on its ongoing establishment of data use routines, but also on faculty's feelings of genuine engagement in decision-making processes.

Belleworth Academy of Arts and Technology hired Ms. Heredia because of her perceived prowess in using data to identify and initiate programmatic responses to student and school needs. Belleworth's hiring process suggests faculty were not only interested in Ms. Heredia's data capabilities, but that they also felt her interpretation and responses to school data were aligned with their own perspectives. As a manifestation of this, since her hire, Ms. Heredia did much work plotting the mission and objectives of the school into a five-year strategy with her ILT. Inherent in this improvement plan was a cultural reorientation of faculty to Belleworth's vision of instruction and the establishment of a common vision of school improvement throughout the school. But changes to power structures and processes of decision-making introduced by Ms. Heredia were challenging. Ms. Heredia advocated for the autonomy of the principal in making some decisions, as well as devolving decision-making authority to councils and committees alongside the ILT. Stripped of their oversight over all decisions made at Belleworth, the ILT became a major source of protest against Ms. Heredia's leadership. As such, while Belleworth had sought a principal who could use data to institute a fresh approach to school improvement, ultimately school leaders were reticent to forfeit some of their decision-making power. The inability of teachers to trust their principal in carrying out Belleworth's strategic plan, reflect on their own instructional practices, and/or facilitate constructively critical dialogue among their colleagues suggests that, even with its intention to rely upon data, Belleworth's leadership team was not necessarily receptive to change. The determination of who

decides what should be done with school data, and in what ways, was equally as influential in Belleworth's decision-making processes as the value of data itself.

Upon entering Woodson as its new principal, Ms. Figueroa was surprised to find that school faculty were initially reluctant to draw connections between student performance data and their own instructional practices. Her own attempts to introduce new systems of data use were met with hostility by a faculty striving to maintain its identity as a teacher-powered school and their own established strategies for improvement. Ms. Figueroa's expression of interest in observing classroom instruction and its association with departmental goals was met with similar animosity – teachers perceived her classroom presence to be more evaluative than observational. It wasn't until Ms. Figueroa was able to constitute personal relationships with individual teachers and find a data use ally in Woodson's research expert that she was able to demonstrate her shared understanding of progress and improvement with Woodson faculty. Ms. Figueroa's relationship with Woodson's lead teachers, in particular, became an essential mechanism in creating momentum for the use of data to understand student progress and improvement.

Although the accessibility and type of data available at each of the three school sites varied considerably, when it came to processes of decision-making, data seemed less a concern than who was charged with making decisions and the ways in which those decisions were made. In the context of these pilot schools, each of which emphasized teacher leadership, how data were integrated into decisions came second to the constitution and maintenance of decision-making systems that affirmed teachers' value as decision makers. Previous research has looked at the positive impact of multiple levels of school leadership (i.e., teachers and administrators) knowledgeable about and committed to data use in decision-making (Feldman & Tung, 2001). The experiences at Belleworth and Woodson College Prep again remind us that the cultural-

political orientation of a school toward data use does not rely solely on individual proponents of data use. At Belleworth, while the principal's reliance on data to inform programmatic and curricular changes was valued by its ILT, the devolution of its decision-making authority to the principal and other teacher committees was a perceived threat to teacher leaders. At Woodson College Prep, faculty regarded their principal's introduction of new data use routines as a threat to the ecosystem of teacher-led decision-making and strategizing already in place. In both cases, and that of The Academy, personal relationships of trust and common understanding among multiple decision-makers were prerequisite for data use. Before allowing data into conversations, teachers and administrators need first to determine the value of their own role in decision-making processes.

## CHAPTER 6

# CULTURES OF “CREDIBLE DATA”

### Introduction

In order to determine how it is that schools use data in their decision-making, it is first necessary to understand what schools identify as “credible data.” This is distinct from simply identifying what data are available for school use. *Data* refer to the full repository of school-based data sources available for school use, a rather exhaustive list (Marsh, et al. [2006] categorize school data types as *input*, *process*, *outcome*, and *satisfaction*). In addition, data writ large carries with it both positive and negative connotations resulting from the unique interactions school stakeholders have had in working with or in being evaluated by data. The term “data,” mentioned in the context of school performance and accountability, thus bears political stigma. In contrast, *credible data* takes into consideration the various perspectives individuals bring in determining what data are practically useful in making school-based decisions, which are relevant to practice and are valid and reliable reflections of student, teacher, and school performance. In understanding how school stakeholders identify credible data, this study seeks to explore the values and priorities individuals place on various data sources. This study hopes to bring some understanding as to how and why certain types of data are incorporated into decision-making processes while others are not.

In the course of this study, the term “data” was most readily interpreted by teachers, principals, and district administrators as a reference to quantifiable student and school performance outcomes. For example, when asked how their school used data, teachers and principals frequently referred to student grades, state test scores, graduation rates, reading levels, suspension rates, or enrollment figures. While these sources do comprise what are defined as

“credible data” in some cases, it was also found that “credible data” also incorporate more qualitative aspects of school activity in both narrative and quantifiable forms. This was true even if such results were not regularly reviewed or required for District reporting. For example, some teachers and principals highlighted the importance of school-administered survey data, including student and parent multiple choice and short answer responses. Still another source of “credible data” frequently cited by teacher, principal and District participants were the affective data regularly collected and used in classrooms and school campuses. For example, the majority of teacher participants emphasized the necessity of getting to know their students on an individual level, suggesting that student background, cultural-environmental context, and work-study habits and behaviors were essential indicators of student academic status and progress. A full list of the types and sources of data raised and referenced by study participants is provided in Table 6.

*Table 6: Data Types and Sources Referenced By Study Participants*

| <b>Data Types</b>                                            |
|--------------------------------------------------------------|
| <b>Student Demographic Data</b>                              |
| Students qualified for Free and Reduced Lunch (Title I)      |
| Percentage of English Language Learner students              |
| Student race/ethnicity                                       |
| Number of parents in student household                       |
| Context of student residence                                 |
| <b>Qualitatively Assessed Student Performance Data</b>       |
| "Anecdotal" information on student achievement and progress  |
| Dialogue & feedback (students, teachers, and administrators) |
| Teacher observation of student behavior                      |
| Classroom observation (student learning and engagement)      |
| Student self-reports of achievement and progress             |
| Student work                                                 |
| Teacher self-report of student achievement and progress      |

Table 6: (continued).

**Quantitatively Assessed Student Performance Data**

Common assessments (school-based alternative to Periodic Assessments)  
 Student attendance rates  
 Student enrollment  
 Graduation Rates  
 Student discipline (e.g., suspension, expulsion, etc.)  
 Student grades

**Standardized Student Assessments (District-facilitated)**

Advanced Placement student assessments (APs)  
 California High School Exit Exam (CAHSEE)  
 Common Core practice test  
 District-facilitated Periodic Assessments  
 California Standards Tests (CSTs)  
 English Language Development Test (CELDT)

**Qualitatively Assessed Teacher Performance Data**

Classroom observation (instruction, classroom management, learning environment)  
 Dialogue & feedback (students, teachers, and administrators)  
 Student reports of teacher/classroom activity  
 Teacher Performance Review

**Quantitatively Assessed Teacher Performance Data**

Teacher Performance Review  
 Value-Added Models of teacher effectiveness

**Principal Performance Data**

Dialogue & feedback (students and teachers)  
 Principal Performance Review

**Stakeholder Satisfaction and Feedback**

Parent survey  
 Student survey  
 Teacher reputation  
 Teacher survey

**Data Sources**

Counselor Records  
 Teacher Records  
 School-based Learning Management System  
 My Integrated Student Information System (MiSiS)  
 Cumulative Files

Within this chapter, case study data are used to show how the determination of what data are considered “credible” as an active conversation among school stakeholders. In particular, The Academy provides a picture of how credible data are identified alongside the construction of teacher and student performance evaluation processes. Examples from Belleworth shed light on what data teachers identify as credible in the course of their daily instruction. Finally, faculty and administration at Woodson discuss tensions over data credibility as they consider their differential applications to instruction and to the measurement of school-wide performance. In some cases, the value of specific data sources is implied, while in others the credibility of data is part of an open debate. The degree to which data are regarded as “meaningful” varies among stakeholders and their unique decision-making objectives. As such, credible data are seen to span across dichotomies that entail formal or informal collection, use in high-stakes public reporting or formatively within individual classrooms, and/or which respond to faculty-led inquiry or District compliance mandates. Importantly, data that are valued as credible within schools are not necessarily systematically collected. Although data collection is traditionally viewed as dependent on an approach guided by methodology and some level of structure in direct response to research or evaluation questions, it is found that school stakeholders draw equally upon systematic and unsystematically collected data in order to inform decisions. What is regarded as credible does not always meet criteria for systematic data. But how a school collectively defines “credible data” is a cultural decision reflective of its values and strategies in improving teacher practice and student learning.

## **The Academy: Data That Defines School Culture**

As a new school, The Academy lacked a longitudinal view of its performance and found itself in a space where strategic planning had taken a backseat to immediate implementation issues. Although it was in receipt of annually-produced District data reports inclusive of enrollment, attendance, and exit exam results, the report represented only a first imprint, i.e., a single image of many required to round out a more resolute picture of The Academy's effect on student achievement. Coupled with the District's suspension of state standardized testing (as it ushers in the Common Core), The Academy stood in the unique position of having little to no historical performance data as a reference for its initial growth and progress. The Academy's experience thus presented an intriguing case of how a school begins to build a sense of how well it is doing. For The Academy, this process started with defining its goals and objectives and implementing its vision and mission. Linking together the disparate pieces of data to which it had access, as well as conducting its own innovative approaches to data collection such as Student Review Panels and its own Teacher Performance Review Program, The Academy found that the collection of data it identified as "credible" was an extension of its values and, in part, what also shaped its culture.

### **Measuring School Vision and Mission**

In discussing his vision for The Academy, Mr. Cooper maintained a diverse portfolio of objectives. Chief among them was the establishment of a defining school culture – a key component of any new school, especially one collocated on the campus of a longstanding comprehensive high school. At the outset of their second academic year, Mr. Cooper described

his hopes for The Academy as “a sanctuary for kids – their home away from home.” To accomplish this, he emphasized:

*We need[ed] to know what our “why” [was]. [You needed to] figure out why you are doing something, then how you are going to do it, then what you are going to do, not the other way around. I would argue that even before “why,” figure out “who” you are doing it for.*

For Mr. Cooper, the identification of this “why” – the purpose for which The Academy stands – was not only essential in distinguishing the character of the school, but the underlying current for all its services and the ways in which it provided them. A cooperative vision of purpose must precede school activity, rather than result as a consequence of such activities. In other words, intention matters. What, then, was The Academy’s intention? Mr. Cooper continued on to explain:

*Skills help to save lives. What’s the alternative for these kids?[...] If everyone believes this, this is the foundation of our culture. If we’re not here and have this to offer, then we are lost and off track.*

A fundamental tenet of The Academy’s approach should be to provide life skills that carry students forward into prosperity and success. Although a seemingly general philosophy, Mr. Cooper made the distinction that character education and opening opportunities for students to engage in “meaningful experiences,” are of even greater priority than plowing through curricular content and advancing students’ human capital through technical skill building. The notion that students “will remember more about how you make them feel more than what you teach” plays a strong role for Mr. Cooper in personalizing education for students. The dark “alternative” to this is an educational experience wherein students lose emotional connection with the notion of achievement.

Gauging the development of school culture, however, is not a straightforward process. Mr. Cooper looked to several factors, some of which were quantitative and others more qualitative. Of these was the total number of students enrolled at The Academy each year. For Mr. Cooper, he believed enrollment figures are one indicator of student engagement. In his words, “The school’s partnership with students isn’t solidified until they’re here.” Although enrollment was lower than expected in The Academy’s opening year (just below 400 students), Mr. Cooper viewed this as a baseline figure for future growth.

Indeed, at the end of its second year of operation, The Academy’s projection for year three was about 500 students. This increase, Mr. Cooper related, would directly translate into additional resources for the school, such as additional teachers and classrooms, and administrative office space on campus. The increase would also stand as evidence affirming The Academy’s overall progress as an institution; as he mentioned:

*The District is looking at us, like oh, your numbers are going through the roof, you're doing great things. In a short amount of time, branding yourself. We're like... in the eyes of the District one of the better pilot schools now, right?*

The culture of The Academy, Mr. Cooper added, is one also measured by attendance rates and grade elevation. These additional pieces of information helped to evidence the effect of Mr. Cooper’s overall message to students: “I DO need to make the effort.” Perhaps, Mr. Cooper suggested, this might also be substantiated by the rate with which work is being turned in, although he acknowledged this would only point to short-term differences in behavior. In contrast, he expected that, in the long-term, a “different feeling on campus, a change in atmosphere” could occur. In fact, as compared with last year, he reported:

*The campus has a calmness and stability to it not seen last year. Students are a little more focused, they're more engaged. A lot of the “drama” has abated.*

*Students are saying to themselves, “I’m here because I have to get my education. I’m taking this institution SERIOUSLY, at home and here.”*

This detection of an atmospheric shift hints at pieces of perceptible information – student engagement, reduced behavioral misconduct, increased student efficacy – but not necessarily in a way that fully demonstrated a methodical inquiry into changes in The Academy’s culture. Although Mr. Cooper made references to several sources of school-based metrics, such as attendance rates and grade elevation, it was not clear that these data sources were systematically analyzed by The Academy. Rather, it seemed to be the general perception and amalgamation of metrics, like attendance and grades, alongside personal observations of student interaction and broader campus behavior that gave Mr. Cooper a read on whether The Academy was progressing toward its cultural vision.

### **Student Review Panels and the Complexity of Evaluating Academic and Behavioral Progress**

Although The Academy had not yet had the opportunity to systematically review data sources linked to its school culture, it is important not to overlook the complexity of measuring such a construct. Mr. Cooper did well to point out that The Academy’s movement toward its goal of a student body engaged in self-efficacious learning was deeply embedded in a process of inspiring both behavioral and academic change within students – distinct, but often confounding factors. Such changes are neither simple to address, nor easy to monitor.

In response to this, The Academy took a new approach to connecting with some of its failing students through Student Review Panels (SRPs). Midway through the semester, teachers clustered by grade or by department and scheduled individual appointments with students they had identified as falling below their potential performance levels. The general purpose of the

panels was to present a collective message to participating students that they were capable of improved achievement and to facilitate plans for progress. A closer look at this activity illustrates how deeply complex invoking changes in student behavior and academic progress can be – assessing improvement in student performance is not a linear endeavor beginning with individual student objectives and ending in measurable personal growth. Additionally, this systematic address of student need exhibited an important process of formally collecting affective data that fed into teachers' comprehensive understanding of student performance.

### Kinsey

Prior to each of Kinsey's appearance before a panel of her teachers, Academy faculty discussed her current state of progress behind closed doors. Although this appeared to be a casual exchange between three colleagues – Mr. Leighton, Ms. Ramsey, and Mr. Easton – the SRP provided a designated time and space for her teachers to compare and contrast their personal evaluations of Kinsey's performance. Mr. Leighton, Kinsey's pre-calculus teacher, first brought to the table his professional opinion: Kinsey does "not apply herself." While observably bright and stocked with potential, her pre-calculus grade was on a swift "fail" trajectory. Mr. Leighton drew on direct observations of Kinsey's classtime behavior in order to construct a picture of her performance – she was often sleeping during class and exhibited very outward displays of disinterest and lack of effort. Ms. Ramsey, Kinsey's drama teacher, was surprised by this assessment, commenting on her observation of Kinsey's energetic demeanor and propensity to answer questions in class. Ms. Ramsey also cited Kinsey's classroom grade (passing) as evidence of her overall performance. She suggested that Kinsey's academic history was strong enough that she "should be college-bound."

To understand why Kinsey might have been exhibiting these differences in behavior across classes, Mr. Leighton oriented the group to Kinsey's perspective. She had expressed not wanting to be in his class in the first place. Instead, she was persuaded into pre-calculus by another teacher. Ms. Ramsey drew upon her own personal experience as a student to provide possible context for Kinsey's apparent lack of ambition. Perhaps, like herself, Kinsey was an "arts-oriented" individual who found great difficulty in motivating herself to engage in pre-calculus. Although Mr. Leighton and Ms. Ramsey acknowledged the importance of Kinsey's personal inclination toward mathematics, this did not outweigh their acknowledgment that Kinsey would need to accomplish the work for pre-calculus as expected by her teacher.

Before Kinsey even entered the room, this small group of teachers had drawn on several data sources in building a common understanding of her current and expected performance: professional opinions, classroom observations, and personal educational experiences, as well as Kinsey's grades, post-high school plans, and self-articulation of motivation. Importantly, Mr. Easton, Kinsey's history teacher, asked Mr. Leighton to explain what "applying herself" would look like and to delineate what Kinsey could do to evidence her improvement. Combined, these pieces of information portrayed Kinsey as a capable student, but one whose behavioral attributes were impeding her academic achievement.

Although faculty had been able to present a snapshot of Kinsey's academic potential within their own circle – as could have been done in a break room chat perhaps – the opportunity to sit down together with Kinsey and communicate with her directly served as another essential method of data collection.

As Kinsey lowered herself into a seat, she was visibly anxious at being confronted by all three teachers at once. Indeed, the panel allowed her teachers to physically represent a cohesive

message showing that they were collectively aware of her current performance. Mr. Leighton began by referencing his twice-weekly discussions with Kinsey about how he felt concerning her classroom behavior and that she was often sleeping during his class. This was echoed by Mr. Easton's impression of her in his history class, placing his face down on his desk and wrapping his arms around the top of his head. Although Kinsey offered up the defense, "I'm tired," Ms. Ramsey gently countered this explanation by mentioning that Kinsey offered up a lot of creativity and energy in her drama class. These statements, while non-consecutive, together showed Kinsey that her teachers knew there was more to her expressed exhaustion. In Kinsey's case, shutting herself out of classroom participation was not acceptable. Eventually, Kinsey voluntarily admitted that she was "just lazy."

Extracting information from Kinsey was a careful process, however, and was not focused on confession. Rather, the panel made a concerted effort to establish a tone of trust, honesty, and constructive criticism in a short period of time. All three teachers were careful to couch their concerns into a positive discussion about Kinsey's academic potential. Kinsey received the message at the outset of the meeting and that it had not been convened to "attack" her. Mr. Leighton and Mr. Easton emphasized her ability to understand the material, even when she is only half paying attention or half awake. Kinsey seemed to internalize these remarks, reflecting back on her success as a math student when she was younger. This external and internal endorsement of her capability allowed the group to pursue deeper questions related to Kinsey's performance.

Ms. Ramsey used some time to confirm Mr. Leighton's earlier claim that Kinsey did not personally want to take pre-calculus. She also delved into Kinsey's post-high school plans in a check against her own "assumption" that Kinsey is college-bound. Alongside Ms. Ramsey's fact-

checking, Mr. Easton probed for additional clues and factors contributing to Kinsey's underperformance. As part of this, he addressed Kinsey's complaint of being tired and asked what time she goes to sleep. When Kinsey explained that she has chores, he asked her to articulate her responsibilities. Although Kinsey only casually mentioned that she babysits for her five-month old niece, Mr. Easton picked up on this detail and asked if Kinsey and the baby live in the same home. Only then was he able to identify a source of her weariness asking whether she had "a lot of late nights with a baby crying at home?" Even in her acknowledgment of this, Kinsey was silent while nodding. For whatever reason, this home life challenge for Kinsey was not something she was quick to bring up herself. It was likely that, had this line of questioning not been pursued, her teachers would not have known quite how to contextualize her fatigue.

But rather than dwell on this complication, Mr. Easton immediately turned to what Kinsey could control in the process of her improvement. Pushing Kinsey in her own admission that she needed to "apply herself," Mr. Easton asked what this means for her, and how as teachers they would be able to monitor and measure this change. Kinsey readily provided indicators of her performance, which echoed almost precisely the words of Mr. Leighton: "I need to turn my work in. I need to read the book. I need to study for the test." Next steps were immediately taken to hold Kinsey to these self-prescribed measures of performance, and a plan for her to make up outstanding work for Mr. Leighton was constructed.

In this brief fifteen-minute conversation, Mr. Easton, Mr. Leighton, and Ms. Ramsey were able to at least partially decode Kinsey's puzzling approach to pre-calculus. Their unified support and concern for her academic well-being set the stage for a line of questioning related to Kinsey's work habits and personal life that provoked honest answers. The three teachers were able to verify what they believed to be true about Kinsey through direct questioning, gaining

additional information as to what might be contributing to her reduced classroom energy, as well as eliciting from Kinsey ways in which both she and her teachers would be able to hold her accountable to her responsibilities as a student. Although any of these teachers might have been able to hold a similar conversation with Kinsey on an individual basis, the panel allowed faculty to pool background information on Kinsey, compare notes, and draw on one another's line of questioning in order to develop a more well-rounded depiction of Kinsey's status and progress. Moving forward, Mr. Easton, Mr. Leighton, and Ms. Ramsey were all apprised as to what Kinsey was expected to accomplish.

Importantly, although a plan forward was charted for Kinsey, the complexity of her scholarship was not resolved. Irrespective of what this meeting did to provide Kinsey with an increased sense of motivation, she will have faced the challenge of persisting through subject content she found personally less interesting, or in balancing what work she needed to accomplish for school and her role and responsibilities at home. For Kinsey, progressing toward her college goal necessarily entailed an intricate mix of academic and behavioral modification. The panel presented an important opportunity to understand Kinsey's classroom performance in context, but it had only begun to surface the complexity of data associated with her progress.

### Adrian

This is also true for other students seen by the panel. There was the case of Adrian who admitted that he doesn't do his homework, but tries, and that he puts honest effort into paying attention during class. Adrian seemed to approach his high school career with genuine intention, but he faced a great deal of challenge in attending school regularly and on time. This was also his second year as a senior. The SRP felt, in Adrian's words, like a type of "interrogation," as his

teachers dug down into the factors influencing his flagging achievement, including his family structure and morning routine. It was discovered that Adrian took a 45-minute bus commute to and from school every day, which he endured as a way of distancing himself from the “trouble” and the people he knew were “trouble” in his neighborhood. Despite Adrian’s expressed desire to earn his diploma and “not hang out with them all day,” Mr. Easton noted that Adrian’s absenteeism and extremely late arrivals to school had large ramifications at The Academy which maintained a block schedule; missing one day of class was really like missing two. Focusing on factors within Adrian’s control, Mr. Easton created an external incentive for Adrian to consistently make it to his class on time: dinner at a restaurant of Adrian’s choice if he could show up on time for the rest of the semester. While Mr. Easton communicated to Adrian that he would be closely tracking his attendance and timeliness over the next six weeks, Adrian would be responsible for monitoring his own work and make-up work responsibilities from The Academy’s online student portal.

#### A Personalized Approach to Systematized Student Background Data Collection

Both of these student examples represent a unique case of challenge and potential. In its first stages of implementation, none of the teachers could have predicted how the SRPs might have influenced student behavior. However, it was clear that this forum had been an important venue for connecting with students on a personal level and for obtaining student background information essential to understanding their classroom performance. While it might have been easy to diagnose sleeping, late, or absent students as lacking “engagement” or “work ethic,” it was much more difficult to pinpoint the root causes and effective supports to improve their disposition. The SRPs thus served as a systematic investigation into those underlying causes,

recognizing that not only would they vary for each individual student, but that they were complicated, interwoven, and, even with this great effort, difficultly detected.

The SRPs serve as an exposition of the type of data relevant to The Academy's objective of cultivating students' self-efficacy in learning. Alongside student background information, teacher observations and comments, the comparison of student performance across classes, and public statements of the types of observable student behaviors indicative of progress are all credible data points. The regular conduct of SRPs could ensure that these data are consistently collected and accessible to teachers and students in the complex pursuit of improved student achievement.

### **Innovations in Measuring Teacher Performance**

Recognizing and exploring the complexity of teaching and learning for The Academy was not just an issue of intimately understanding student perspectives. As part of building a school culture wherein students take responsibility for their own education, there was also a vision of teachers who create the fabric of a supportive, effective learning space. While not an uncommon expectation within schools, The Academy strove to incorporate this vision into a unique approach to teacher performance reviews.

The Teacher Review Program (TRP) recently instituted at The Academy was imported by Mr. Cooper from his former campus. The Academy's stated objective in introducing the TRP was "to establish a professional Peer Collaboration policy and Teacher Evaluation process that effectively promotes and maintains a dynamic and high qualified teaching staff who actively support the mission of the school" ("The Academy's Teacher Review Program Overview," 2013). In opposition to the conventional notion that the purpose of teacher evaluations is to

“weed out bad teachers,” The Academy emphasized supporting teacher professional development through peer review. The underlying rationale for this structure acknowledged teachers as autonomous professionals best-placed to review teacher professional practices.

The teacher review process begins by asking the teacher under review to reflect on specific expectations and standards of performance organized into three general areas of assessment criteria: their classroom (i.e., knowledge of effective practices, use of student-centered instruction, or responsiveness to student needs), The Academy’s community (i.e., participation in group or school-wide activities, communication with parents, students, and faculty members, or supporting and integrating character education concepts), and their vision and goals (i.e., maintaining currency in subject matter and profession, willingness to expand technology use, or implementing professional development learning in the classroom). In addition, each teacher is asked to select five areas of focus from the District’s “Teaching and Learning Framework Focus Elements,” which peer observations would hone in on. The overarching standards organizing these elements included: planning and preparation, classroom environment, delivery of instruction, and professional growth. Reviewed teachers were advised to choose areas of focus perceived to be of most benefit to their own professional practice (“Teacher Development Focus Area Worksheet,” 2013). Lastly, teachers under review were asked to administer student surveys to their classes the academic year prior to their review. These were developed, administered, and initially reviewed by the teacher.

Over the course of several months, teachers under review met with their department and a Teacher Review Committee to engage in genuine discussion about their progress along the expectations and standards of performance. Alongside these meetings as evidence of class performance, their departmental colleagues compiled student outcome data such as test results or

grades. Teacher teams then discussed the strengths and weaknesses of the teacher under review, as well as their overall impression of the teacher, providing their feedback to the Committee.

The Committee ultimately took into consideration departmental feedback, student survey results, professional portfolios or interviews, administrative contributions, and data gathered from classroom observation results focusing on the five target areas identified by the teacher under review. Upon developing their final recommendation, the committee met with the teacher under review to provide feedback and an opportunity for discussion (particularly with respect to classroom observation findings and comments). Final recommendations were ultimately passed on to the principal, who was charged with making a decision on the status of all reviewed teachers.

Underlying this year-long process was a focus on teacher support and development that aimed to enhance collegiality and professionalism among The Academy's faculty. It was based on the premise that teachers are "as effective, if not more, than administration to consult and evaluate fellow teachers" ("The Academy's Teacher Review Program," 2013). Through self-reflection, observations, the sharing of ideas and skills, and the consultation of teacher in improving their practices, The Academy's TRP was seen as a way of promoting meaningful dialogue, greater transparency, and useful feedback among its faculty, complementing its vision as an egalitarian and consensus-driven school.

Performance review systems supportive of reflective teacher practice, participants at The Academy argued, require teacher-identified data sources deemed reliable and relevant to their own instruction. The Academy's TRP thus drew on a host of different data sources to construct a holistic picture of teacher performance and practice. While standardized test scores, student grades, professional portfolios, classroom observations and feedback from colleagues and

administrators are not uncommon components of teacher evaluations elsewhere, what distinguishes The Academy's collection of data is the perspective and orientation from which the data are derived.

The Academy made clear that the assessment criteria guiding classroom observations are not externally imposed "checklists" of activities and elements either present or not present at the time of observation. Rather, the teacher under review is expected to choose areas of focus on which they desire guidance from their colleagues. In this way, the feeling of being "attacked" by someone searching for an endless inventory of teaching qualities and practices is replaced by a targeted, thoughtful approach to improvement generated from a teacher's expressed needs. The emphasis on teacher self-reflection is considered paramount to an effective process of review and stands as a central pillar of the evaluations. Much like the substance of the SRPs, a teacher's honest acknowledgement of strengths and areas of improvement is seen to promote deeper conversations about constructive self-progress.

Student surveys were also considered a crucial component of teacher feedback and reflection. No specific guidelines were given to teachers as to which classes they should survey or the content of their surveys. Because of this lack of external oversight, teachers ultimately determined which results were brought to their colleagues. Irrespective of what was eventually shown to the departmental teams and Teacher Review Committee, this piece was seen as one of the most insightful perspectives of a teacher's practice for their own personal consideration. Mr. Cooper reflected on his own experience as a teacher reviewing this form of student feedback:

*To me there is nothing more valuable than kids telling you what they think about you. Anonymously. Right? And there's always going to be one or two, that's like so powerful that it hurts. But the vast majority, or the average, whatever they're*

*telling you? That's where they're at. That's their perception about the teacher. And I honor that. And I value it. It's like... to me it's gold.*

For Mr. Cooper, there was no escaping the honesty of students. Although he made reference to the potential brutality of some comments, there was no hiding from the general class perspective. For Mr. Cooper, student surveys offer an unparalleled level of accuracy worthy of serious consideration. Mr. Easton, having been part of the same faculty as Mr. Cooper at their former school, it was also helpful to usher the TRP into The Academy. His estimation of the importance of student surveys echoed that of Mr. Cooper:

*And I think that to me, that's probably the most important voice in all of this, is the students.... You know it's interesting, because I do it and I get these back, and for the most part, students really liked me as a teacher. But there are some times that they say things that are so spot on, but they HURT, you know? Because they're so spot on [smiling]. And you go like, but... but... But you can't argue, because they're sitting there telling you yeah, you don't do this very well, you know? And you have to kinda' look at that and say, yeah, it's really true, I DON'T do that very well. You know? And that's something you don't get in a normal sort of review.*

Like Mr. Cooper, Mr. Easton recalled the fierce honesty with which students offered their comments. Both educators even emphasized the physical pain associated with some of the more biting remarks that “hurt.” Mr. Easton conceded that while most of his students liked him as a teacher, he was unable to avoid intermittently harsh feedback. As defensive as he might have felt about what was said, at the end of the day, he succumbed to the realization that these comments were agonizing because they were true. For these reasons, Mr. Easton promoted the student voice as crucial to a teacher’s genuine reflection on his or her professional practice.

Another unique feature of The Academy’s TRP was its treatment of accountability. The idea that one is reviewed by peers rather than by administration – who may or may not have a strong sense of what a teacher’s day-to-day practice entails– instills a sense of validity to the

process of evaluation. As a result, those who have reviewed their peers' performance come from a place of fair judgment and professional understanding. If the review is implemented more as a guided discussion or even as mentoring sessions, it becomes still more personalized.

Buttressing this fluid structure of dialogue and discussion of practice were the standards and expectations detailed by the TRP, as well as the "focus elements" from the District's Teaching and Learning Framework, which outlines LAUSD's expectations and standards for "effective teaching" and associated exemplary practices (Los Angeles Unified School District, 2013). These guidelines added an important sense of formality and frame to the review, as did the systematic process of documenting findings and recommendations by department teams and the Teacher Review Committee. Ultimately, however, the feeling of accountability seemed to be one that is internally driven by The Academy and was oriented toward the support and growth of teachers. As another teacher, Mr. Knowles, put it:

*Our program, our Teacher Review Program at our school that we've developed on our own, we're... one of the only, if not THE only school in the district that has decided to opt out of the District Teacher Growth and Development cycle. Ours is designed to help you become a better teacher. Theirs is designed to identify teachers to get rid of. So ours is CONstructive, and theirs is DEstructive.*

The notion of truly improving teacher practice, from his perspective, demands that the process of teacher evaluation re-think its orientation:

*Where is the "We want to make you better before we determine if maybe you shouldn't be here?" A lot of these competing systems are like, okay, there's a black mark against you, you get so many black marks, you're gone.*

For Mr. Knowles, The Academy's TRP was unique in distancing itself from the District's standard process of teacher review and evaluation. He viewed the latter system to be primarily punitive wherein teachers were subject to dismissal based on the accumulation of "black marks"

or documented areas of “non-performance.” In contrast, the TRP at The Academy was built upon a sense of accountability to one another as teacher professionals rather than to “higher-ups” distanced from the classroom or anonymous “District” decision-makers. In The Academy’s system, Mr. Knowles emphasized the importance of allowing teachers the time and space to correct and improve upon their practice.

Data collection under the TRP occurs over the course of an entire year and is intended to result in a thorough reflection of teacher practice through peer mentoring and deliberative dialogue. Because the process relies so heavily on teachers to govern the process and appraise their colleagues’ performance, however, a great deal of trust is placed on faculty to compose an evaluation that is objective, fair, and constructively critical – Academy faculty are the stewards for teacher accountability. The type of data incorporated into teacher performance reviews, and the ways in which classroom data are collected, analyzed, and integrated into performance appraisals are seen as a manifestation of The Academy’s confidence in teacher voice, authority, and professional judgment.

### Challenges to Implementation

However, design of the TRP and its actual implementation at The Academy were quite different things. The ability to rely on all faculty members to authentically engage in constructive dialogue around not only teacher practice but also peer-reviewed performance required careful cultivation. As a result, the kind of teacher performance data The Academy would have liked to gather faced the challenge of reorienting staff perspective away from bureaucratic systems of accountability and toward a data gathering process that actively involved individual faculty members.

Charged with leading the TRP within The Academy, Mr. Easton described its implementation that year as “interrupted” and, in some places, “superficial.” Although he recognized that adaptation was, in some ways, a process of trial and error, Mr. Easton felt that the TRP “definitely [hadn’t] been as deep or as thorough as anyone involved in the process, especially Mr. Cooper and me, [wanted] to see.”

Of course the notion of time was a contributing factor. Mr. Easton described that nothing about The Academy was yet running on “auto pilot,” and its “hard-working and very well-intentioned” teachers struggled with making room in their schedules for the review process alongside other substantial commitments. However, Mr. Easton emphasized that the absence of time also impacted “that delicate nature that comes with being able to discuss what we do in the classroom.”

He went on to explain the feeling of “exposure” that accompanies participation in the TRP. Teachers’ propensity toward “protectiveness” and a defensiveness around their classroom performance, in Mr. Easton’s opinion, stem from the customary “isolation” in which teachers work, as well the unpredictable nature of their everyday instruction. He expanded:

*There's so much that teachers do that's out of our control, you know? You can prepare a lesson, but if Jose or Vyas is going to have a bad day, they can sink that lesson in a matter of three minutes. And so, with all of [those] external forces that can really change the way that our classroom goes on any particular day, it's pretty nerve-racking to think that someone's going to come in and watch you teach.*

Here, Mr. Easton raised an issue echoed by several other study participants: the nature and number of factors influencing the implementation and effect of a lesson are in many ways uncontrollable. For even a seasoned teacher like Mr. Easton, who had maintained 27 years of teaching experience and skill, preparation is only part of the equation for a successful lesson.

Rather, “external forces” – e.g., students’ emotional states and behavior – come into significant play in delivering that well-prepared lesson. This lack of control and predictability is compounded by the rushed pace of a school under development. Teachers’ meaningful reflection on practice is, in effect, frustrated by a sense of incalculable instructional effect and impeded by days overflowing with various administrative and committee responsibilities.

The peer evaluation of teacher performance thus relies on a foundation of mutual trust, something that takes time to build among colleagues. Trust is required to both accept what cannot be controlled and to believe that others will justly consider such variables in the valuation of one’s performance. As Mr. Easton described it, trust between teachers is an aspect of the TRP that is simultaneously required and generated by the review process. One of the challenges in constructing this sense of mutual rapport was breaking The Academy’s teachers out of their conventional evaluative paradigm:

*And so building a sense of TRUST that a program like this is meant to engender is a challenge, you know? Teachers are just used to being evaluated. You know, tell me I did a good job, I'll choose a lesson that you can come watch, which has got all the things that people are looking for as opposed to being sort of... genuine.*

The culture of performance and accountability encompassing conventional systems of teacher evaluation, Mr. Easton argued, propels teachers through a series of scripted activities simply meant to meet evaluation requirements. In this theater-like production, teachers hand-select lessons containing mandatory elements of “good teaching,” and in return, expect to be told they are doing “a good job.” The exchange is one that lacks ingenuity and concurrently strips the teacher of a sense of ownership over the findings.

Providing an example of how this type of conditioned mentality has played out in the year’s implementation of the TRP, Mr. Easton explained:

*We had a teacher just this year who I went to observe, and when we were debriefing [his lesson] as a group of teachers, we got into a larger discussion about student plagiarism... What is plagiarism? And how do we teach kids not to do it? And what does it look like? And what we do as teachers to try and avoid it? And I actually thought it was a really good discussion, and actually kind of an important one. And it's not one that would've come up right away.*

*But about, I don't know, maybe about six minutes into the discussion... and it was a little contentious at parts – teachers are disagreeing with each other, and to me that was GREAT. You know? It was like okay, having real discussion about something, and people are buying into their sides emotionally, and they're invested in it. And it's these kinds of the discussions that we wouldn't normally have under a different sort of process.*

*And one of the teachers, that was actually the one that we were observing, said, "Hey I don't know if we're supposed to be talking about this. I thought we were supposed to be talking about how great a teacher I am."*

*And I'm like, no! This is EXACTLY what we're supposed to be talking about, you know? If this comes up in the discussion of us observing your classroom, that's a good thing. Because it's what we're supposed to be about – observing each other's practices. What do we get out of it? And what struggles do we have? It's not just about, "Hey, you're great, here's a lollipop. Come back in another six months and make sure you floss." It's really supposed to be about those discussions that build us as a community and to help us understand each other. Help us, sort of, identify the struggles that we have that are in common, and things that make us different than a larger school.*

In this scenario, Mr. Easton described how one of the review meetings transpired into an in-depth discussion of student plagiarism and the many nuanced questions surrounding its identification and treatment in the classroom. For Mr. Easton, this dialogue was important not only because of the topic area, but also because teachers were engaged in a deep debate that would ultimately “build [them] as a community and help [them] understand each other.” Mr. Easton saw this being accomplished in the process of “identify[ing] the struggles that [The Academy’s teachers] have that are in common.”

Rather than to encourage this exchange, however, Mr. Easton points out that the teacher under review halted the process, not only questioning the appropriateness of the conversation in the context of the review, but also (perhaps jokingly) asserting that the meeting's real purpose was to offer him praise. Mr. Easton attributed this expectation of platitude (however sincere) to a culture of superficial accountability checks similar to those encountered over teeth cleanings at the dentist. This mindset, he suggested, is derived from years spent working within the District's systems.

While frustrated by conventional perspectives of teacher evaluation, Mr. Easton knew that breaking free of this paradigm is a matter of capacity building. He comments,

*It doesn't just happen overnight that [teachers] go from being like, "I'm part of this huge bureaucracy that tells me what to do and how to do it. I don't really have to think too deeply about it because that's not my job – I can just go and teach," to being in a situation where we're accountable to each other. Where the decisions have to be made by US. And you know, I don't think that we're there yet as a school.*

The "cultural conditioning" that Mr. Easton highlights is one that displaces teachers from a seat of genuine accountability. Mr. Easton saw the District-directed teacher as one automated to simply follow instructions rather than to meaningfully engage in activities like the TRP. In Mr. Easton's eyes, persistent submission to District accountability mandates eventually devolves into a teacher-adopted mantra of "that's not my job" when it comes to "thinking deeply." Such a loss of autonomy and sense of control among faculty is one of the greatest challenges Mr. Easton faces in implementing the TRP, which demands that teachers take a central role in decision-making and professional feedback. The type of cultural revolution required to build an effective TRP is not something that will happen "overnight."

Indeed, as Mr. Cooper confirmed, honestly confronting one's practice and then sharing it with a group is not an easy endeavor. The active determination of what data should be considered in one's own performance review puts faculty in the position of reconciling the resulting outcomes— positive or negative. Mr. Cooper believed this “fear” is probably the cause for difficulty experienced last year in conducting student surveys:

*There were about two or three teachers last year who didn't do the surveys. And I was like, drilling them. I sent reminders. They [knew] they had to do it. They chose not to, because they were afraid to see what the kids are going to say.*

While The Academy's approach to teacher evaluation was one that de-emphasizes the “high stakes” threat of punitive measures, Mr. Cooper highlighted trepidation from The Academy's faculty to brave genuine feedback. Approaching one's practice with openness and honesty is dependent upon a generally cohesive community. While such a culture cannot be forced, it is something that Mr. Cooper believed could be actively nurtured. Faculty must learn to not only be vulnerable with one's peers and accepting of potential criticism, but that it is also important to be coached in providing constructive feedback. Mr. Cooper provided an example:

*You know, to be vulnerable means that I trust you that you're going to say things that I may not want to hear about myself. But can you say it in an objective kind of way without it being personalized? You know, I'll talk to the staff about witnessing before you react.... You know, just like take it in--don't reject it, don't take it personally. But when you give feedback, or when you say something to somebody, it's like... “I noticed you were talking in class, and giving directions to the students in class they weren't engaged and paying attention. And it seemed to me the way you were... like you chose not to address that, and I was wondering why you did that?” Without any judgment.*

This kind of skill building requires time and attention, and Mr. Cooper acknowledged more resources would need to be allocated to capacity-building. At the time of his interview, he

saw that some of the teachers participating in the TRP regarded the process as “perfunctory,” treating it more as “something to get through” than a meaningful exchange.

The Academy had not yet established the systems and architecture required to routinely analyze standardized measures of student, teacher, and school performance. However, the kinds of data it chose to prioritize, such as the data collected by way of its SRPs and TRP, reinforced its overarching mission and vision. The credibility of data stemmed directly from the school’s organizational values. In order to determine how students might better direct their own learning, for example, teachers collaboratively investigated the contextual, background, and motivational influences behind students’ academic and behavioral performance. As a way of closely understanding teacher practice, The Academy’s evaluation of teacher performance was guided by teachers’ self-reflection, peer mentoring and discussion, and student feedback. The culture of the school thus took on a reciprocal relationship with the data it chose to collect and the ways in which they were collected and analyzed. Establishing conventions of trust and mutual accountability between teachers, between teachers and administration, and between teachers and students will be an ongoing quest for The Academy as it continues to define what data are most credible in the assessment of its growth, achievements, and challenges.

### **Belleworth School of Arts and Technology: Acknowledging Current Teacher Data Practices**

Belleworth School of Arts and Technology looked forward to transforming its data use practices under the guidance of new administration in the coming year. Ms. Heredia, its new principal, had been working with faculty to routinely review student performance data as a way of guiding decisions around the development and implementation of student support programming. Although teacher leaders within Belleworth were enthusiastic about using school-

based data in these new ways, it must be recognized that, for several years preceding, faculty had been relying on their own individual data use practices. A look at what data teachers depended upon to inform their own estimation of student achievement and progress reveals several classroom-based data sources conventionally overlooked by proponents of more standardized outcomes.

Much like their colleagues at The Academy, teachers at Belleworth relied heavily on student background and contextual data. While not always systematically collected and reviewed, such data were nevertheless considered credible sources of student information and critical to strong instruction. In this section, examples provided from Belleworth show how some teachers explicitly employed these data within their instructional approaches to encouraging student participation and student engagement with curricular content. Additionally, one teacher walked through what District-collected data he extracted from students' cumulative academic files in order to complement his experiential understanding of student performance and to inform his pedagogical strategy.

### **Community-Based Intelligence**

Mr. Nuñez had been teaching for a total of five years and had been at Belleworth for three. Although he saw himself as “early in his career,” he took pride in his longstanding relationship with the school’s surrounding community. Having moved to the neighborhood when he was four, Mr. Nuñez was committed to working within his local area for the foreseeable future. For Mr. Nuñez, his geographical connection with the students at Belleworth was a key component to his success as a teacher:

*I guess you could say that I connect better with the students. I can UNDERSTAND them better, which in TURN gives me a better environment and a*

*better peace of mind in MY classroom. Because I understand where they're coming from.*

Being able to place his students in exact locations around the neighborhood, as well as having an intrinsic understanding of the context in which they lived, gave Mr. Nuñez a unique point of access to his students. His ability to “understand where [his students] are coming from,” was valuable information in deciphering what they bring with them to the classroom. This affords him not only a personal “peace of mind” but also allows him to foster a “better classroom environment.”

Mr. Nuñez went on to explain that his knowledge of local landmarks and slang words enable him to surprise students with his “insider” understanding of their own day-to-day contexts. Students’ reactions of “How the heck do *you* know?!” are evidence to Mr. Nuñez of a “gateway to building rapport with them.” And building rapport with students, Mr. Nuñez emphasized, is an essential driver of both teacher effectiveness and student achievement:

*That's KEY. If students don't make that... connection with you, that is not connected to the content that you're teaching in the classroom, it's going to be tough. You know, you won't get them to respond, or... they won't open up. You won't understand how to help them. You know, they could easily just shut down.*

From Mr. Nuñez’s perspective, students’ propensity to make connections with the math content he teaches is dependent upon their connections with him as a teacher. Whether and how they offer responses in class, or the degree to which they feel comfortable exposing their own vulnerabilities as learners is based not only on the quality of instructional delivery, but also on their personal interaction with Mr. Nuñez. Without this bond, a student would be very likely to “just shut down,” refuse to engage with his or her teacher, and present a much more difficult challenge in determining what steps need to be taken to encourage his/her success.

Mr. Nuñez offered the example of a student in his math class, Gabe, who was struggling with the material at the beginning of the year. “He was always just there... Nothing ever... He never, he seemed not to be engaged,” Mr. Nuñez described. However, Mr. Nuñez soon discovered he could make a personal connection with Gabe, whose older brother he knew from high school. Pulling Gabe over one day, Mr. Nuñez explained that he knew where he lived, that he knew Gabe’s older brother growing up, and even suggested some specific issues with which he might be dealing. He recounted:

*And little by little... there came a point where he actually started... he broke down and cried. “Oh you know, I’m very emotional, I just... it’s just hard for me to express myself, when I DO I feel bad!”*

*Ever since that day, you know, every day he had a question.... Ever since after that, he would ask... MINIMUM three questions per classroom, to a point where the students got used to it. So they were at first... “Where does this guy come from?... Now he’s asking all these questions. Like, who is he? What happened to him?”*

In this particular case, Mr. Nuñez was able to leverage his personal knowledge of Gabe’s background and home life to establish a line of emotional trust. It was only after he was able to broker this more intimate relationship that Mr. Nuñez noticed an increase in Gabe’s classroom engagement exhibited by regular questioning. Mr. Nuñez highlighted that Gabe’s sudden shift in classroom behavior was even recognized by his peers – an indication of Gabe’s dramatic swing from being a low-profile student to a regular voice in the classroom.

Even with Mr. Nuñez’s long-standing connection with the local community, he didn’t have personal ties with every student as he did with Gabe. When it came to gathering background information on other students, Mr. Nuñez looked to a number of different sources:

*Like sometimes when we’re going through attendance or something sprouts where we notice a CHANGE in a student... we peek into their file. You know, what I see*

*besides their past history is where they live. You know? Where do they live? Okay so, X student lives here. Oh I remember, that's the neighborhood where certain people hang around, and certain people do certain activities. So maybe that's why that student is acting this way. Or maybe that's why he started to be different, or speak differently.*

Because of his own personal grounding in the community, Mr. Nuñez attributed much of his knowledge about students to his ability to place them in their local context. Knowing a student's past academic history is further enriched by knowing where they live. An address, for Mr. Nuñez, is more than a geographic orientation. Rather, an address provides clues as to a student's neighborhood culture and the activities and people who surround them. For students in which Mr. Nuñez notices a sudden change in behavior, this kind information can help to contextualize why a student might be acting differently, taking on different personas, or even speaking differently. These pieces contribute to Mr. Nuñez's more comprehensive awareness of "this is what he brings to the classroom every day," and his ability to say to a student with grounding sincerity, "I understand where you're at. This is what's going on here."

As much as a teacher may be able to demonstrate his or her knowledge of a student's personal orientation, however, a strong teacher-student relationship is equally reliant upon student buy-in. Mr. Nuñez was able to explain how these connections are made when, for example, he is able to exhibit for his students knowledge of their background beyond just their address, their phone number, or "the extra stuff" that is not on official record:

*So that alone says, "Wait a minute, this guy KNOWS something a little bit. Well how? Why? You know, he's telling me this, he's telling me THIS, well he knows THIS... "Now the advice that I'm giving them, well... "I BELIEVE him, because he's telling me all of these things that I don't expect from ANYONE. How does he know?"*

*... I guess our conversations have a little bit more, they're more valuable. They have some concrete validity based on those initial conversations that we have,*

*you know? And that in TURN, once they're in the classroom, they see me different. "Oh, this guy knows what's up. This guy knows my block."*

Mr. Nuñez believed he was able to garner a substantial level of trust with his students because he could evidence locally-relevant knowledge. Students valued this, not only as a matter of feeling “understood,” but also as an internal check of their teacher’s integrity and legitimacy. For Mr. Nuñez, the ability to exhibit community-based intelligence was a matter of establishing his validity as a teacher and mentor. His advice, as a result, was given weighted consideration by students, and he is acknowledged as a respectable authority – a figure of knowledge in the classroom. Such regard, suggested Mr. Nuñez, garners an atmosphere of mutual respect, and respect for learning in his own classroom environment.

Although Mr. Nuñez’s emphasis on building rapport with his students may initially seem to be an auxiliary topic to “data-based decision-making in schools,” what he is able to so succinctly articulate is the way in which personal student backgrounds are key data in facilitating his instruction. Mr. Nuñez leveraged both his own knowledge of community culture as well as small clues and details about his students’ personal lives in establishing a foundation of trust with his students. This, in turn, translated into a positive working relationship and one that more effectively engaged students in curricular content.

### **Building Student Rapport as a Means of Identifying Learning Strengths and Needs**

Because of his long-standing involvement in the community in which he teaches, Mr. Nuñez may be seen as an outlying example of a teacher who integrates students’ cultural context into instruction. At least seven other teachers and principals within this study, however, explicitly mentioned the importance of rapport-building with their students – and getting to

know students on an individual level – in delivering effective, substantive, and meaningful instruction tailored to students’ specific strengths and needs.

To highlight another example from Belleworth, Ms. Gavin explained the kinds of data she collects from students at the beginning of the school year as a locus for her instructional approach which relies heavily on group work and student-to-student interaction:

*OK, so, beginning of the year, first thing I do, I let the kids sit wherever they want on the first day of school. I have them fill out a similar card to these index cards. They put their name, they put their phone number, they put their address. Then on the back of it I ask them three questions. I say, like, tell me three things that you like to do. So that’s the first thing I do. That gets me to know the kids and what their likes and dislikes are. As a general for the class.*

*Because I let them sit wherever they want, what I’m doing is, I’m trying to see who their friends are, who they’re going to talk to, and who influences them. That’s the biggest thing. Once I get to know... the individuals in the class and as a whole, then I start figuring out my strategies that I use per class.*

*In general, any teacher should know an individual student because you need to know what their weaknesses and their strengths are. AND you know need to know how open they are to working with other people in the class.*

*So for seating arrangements, you might not know that this kid doesn’t like that one because maybe they dated two years ago. Everybody in the class knows, so when you DO start moving them around and you start hearing the [gasps] when you move them, you think like, ok, something happened. I gotta’ figure out what it was so that I know what the dynamic was. With that, too, I always make sure that if there was a student that REALLY is adamant about I don’t want to do this, I don’t want to do this, you need to figure out why. They always have to have a reason.*

At the classroom-level, Ms. Gavin detailed how she uses both personal information about a student – for example, their purported likes and dislikes – as well as her observation of students’ in-class interactions and relationships to determine seating arrangements and to compose cooperative working groups. Ms. Gavin submitted that her instructional strategies are

even responsive to her students' classroom relationships – who they talk to and who influences them. Ms. Gavin emphasized that, at the individual-level, the value of this information in understanding how “open” students are to working with other people in the class, as well as in uncovering students' strengths and weaknesses, is important. She was determined to consider each student's personal orientation to learning as a component of the entire classroom dynamic. Rather than base classroom moves solely on her own read of student perspective, she encourages her students to actively participate in supplying information as to reasons why they may or may not be inclined toward certain classroom activities.

Ms. Gavin continued to explain how this type of student-level information factors into specific instructional strategies and not just her organization of the classroom. Ms. Gavin's knowledge of students as individuals allows her to comprise working groups that balance out their strengths and weaknesses. This becomes imperative to ensuring that the varying roles and responsibilities within each group are well-represented and enables her lessons – structured around the groups – to move forward. Ms. Gavin explained how her own instructional strategies appear different from class-to-class as she adapts to the natural skill variations between them. For example, while she was able to give her first period class its own space and time to move through the material, she found that her fourth period class required more structured time and accountability measures to ensure that they completed all required tasks. She noted that her first period class was her “highest achieving” class, and later described her fourth period class as her “lowest achieving.” But beyond simply categorizing her classes in this way, Ms. Gavin made conscious instructional moves to attend to these differences. She admitted that this sort of flexibility is time and effort intensive. When asked if she could “read” the class after just a few days, she replied:

*It's a lot of things. No, it takes a while. It takes a while. It's... first of all, planning. You know like, you've met me to know that I plan. Like, I'm not a, "Let's wing it," type person. I plan a whole lesson, and I always plan a lesson for like, the highest achieving class, because I have very high expectations. AND I'm really good at changing throughout the day. So, like, if something didn't work in first period, I'm already able to be like, OK, next period I already know what to do.*

*So basically, I plan this lesson to be at my highest achieving class. Once I've done that, and if I see that they're not getting it, or it didn't work, after about a month-and-a-half of school, I can figure out WHY isn't it working for this one period and why is it for...*

*And that's one thing a lot of teachers DON'T, don't do. They say, like, "You know, in my second period, I was able to do this, but my fourth period couldn't do it." And I say to them, "Why do you think they couldn't do it? Was it YOU, do you change the way you facilitated it? Is it the kids? Are they not understanding? What are you doing differently in that class?"*

*So first you have to look at yourself. Did I do anything differently from class to class? And then from there, then I know, OK, this class needs more sentence starters. For some reason this class just is NOT as... I don't know, maybe in science they're not capturing the image or something, so they might need something different.*

Ms. Gavin acknowledged the intention and effort required to facilitate adaptive instruction. She begins with careful planning and calibrates her lessons to her own idea of what a "high achieving" class will be able to accomplish. As her lessons are being implemented, however, Ms. Gavin is constantly tinkering with aspects that need to be altered or improved to fit the specific needs of each class. While her colleagues sometimes struggle to understand why a particular lesson may have worked with one class and not another, Ms. Gavin is quick to point out the need to think concertedly about the myriad of factors influencing each distinct group of students. This entails some self-critique on the part of the teacher. Reflecting on the differential effects of a lesson across classes, Ms. Gavin is as bold in her approach to self-examination as a

source of data as she is to consulting her toolbox of various instructional strategies in the attempt to effectively reach her students.

### **Student Data as Contributive to, Rather Than Predictive of Achievement**

Both Mr. Nuñez and Ms. Gavin provide examples of data collection that is in-depth, personalized, and attuned to individual student needs. However, the credibility of data gathered from students and classrooms in this way, as seen in the case of The Academy, rely on teachers' ability to carefully and considerately extract personal data in a way that both respects and propels respect for students' understanding of classroom content. What is distinctive about Mr. Nuñez's and Ms. Gavin's approach is their process of using elements of student background as a springboard for further dialogue, questioning, and probing into a student's individual motivational and behavioral orientation. This stands in stark contrast to the propensity of some teachers to create general assumptions of students based on their past teaching experiences and aggregate patterns in student demographics. For example, one teacher at The Academy noted, "We have a body of students who have poor work ethic and no culture of studying." He theorized that "because this is an arts-focused school, the kids have this aversion to math," and went on to say that "middle, upper-middle class white and Asian kids – it doesn't matter what teacher they have, they're going to do well.... And because their parents are all very wealthy and educated, these kids kinda' look down on their teachers."

Another teacher at Belleworth observed:

*My typical student here is, I guess I want to say they're pretty lethargic about learning.... I mean, you can poke and prod, you can get them to do work but a lot of times it's not quality work.... A lot of them, I think there's more going on in their world than what they let on. That's why we get some... you know, a lot of the kids are just not... they're very passive about their education.*

He saw a number of factors contributing to this “typical” student stance, including a lack of parental involvement:

*A lot of [students] have a lot of home issues. You know, whether it's economic, whether it's social, whether it's just not... You know, your parents are working, they're never really around. You're basically raising yourself.... Yeah, they have parents there, their parents are around and they're either too tired to really do much, or you know, they're just really working. I think it's hard when you're a young person and you're taking on these responsibilities because you don't have anybody there to guide you through that.*

The point here is not to single out these teachers and critique their belief systems or the professional opinion to which they are entitled after many years of work in the classroom, nor is it to say that either of these teachers lacks empathy for their students or their personal challenges. Rather, these examples serve to exhibit different ways in which contextual student information may be interpreted and used. In these two discrete examples, patterns in student race and ethnicity, socio-economic environment, content preferences, and parental involvement are some of the many factors these teachers see as interacting with students' motivation and sense of responsibility. However, in these two cases, such variables seem to be regarded either as self-explanatory or predictive of student behavior. This is distinguished from the use of student background information as a starting point for deciphering individual student's engagement in the classroom, followed by teacher adaptation and modified instruction.

### **Teacher Interpretation of Student Statistics**

It is important to note that in-class student data collection, such as that described by Mr. Nuñez and Ms. Gavin, are also complemented by teachers' use of more standard metrics of student performance as a way of detecting student needs and strengths. Alongside his wealth of community-based knowledge, Mr. Nuñez is also a strong proponent of conducting background

research on students by reviewing their cumulative files. Termed “cume files” or “cumes” by many teachers, these District-maintained records are a longitudinal compilation of a student’s transcripts, assessment scores, teachers’ comments and evaluations, and student work samples, as well as school registration records, such as vaccination records, birth certificate, and, where applicable, documentation of immigration throughout his/her school career within the District. All of this information is amassed into a physical file designed to travel with each student to their school site(s). Teachers and school administrators have direct access to these files, and while parents and students have rights to their own records upon request, Mr. Nuñez admitted that few know they even exist.

When I met with Mr. Nuñez for our second interview, I found him sitting in his classroom at a cluster of student desks with two thick manila folders stacked next to his right arm. Having mentioned in our first meeting that he was trained by his teacher certification program to refer to and regularly review cume files, Mr. Nuñez kindly offered to give me an introductory tutorial on their analysis, something he attempts to do regularly with incoming teachers at Belleworth.

*We first plodded through the file for Jeremy. Mr. Nuñez cracked open the file and began with the records adhered to the inside cover of the folder – sticker printouts of Jeremy’s transcripts. First we saw Jeremy’s grades and credits from his first semester at Belleworth, as well as those from middle school. These, Mr. Nuñez suggested, allow us to get a “glimpse” of the student’s academic progression by subject. Alongside each of Jeremy’s class letter grades are additional letters, coded to represent “excellent,” “satisfactory,” or “unsatisfactory” in categories like “work habits” and “cooperation.” Glancing through Jeremy’s transcript, Mr. Nuñez posed a scenario: “So if I’m a science teacher... okay, well, why didn’t he do so good in my class? Well, let me check his past science classes.”*

*Running his finger down the list of grades and classes, Mr. Nuñez spotted a small hiccup in Jeremy's grading pattern. "Well that kind of explains it. Go back to seventh grade. Well, I see seventh-grade there's a..." he pointed to a "D." "Where's the first semester? Interesting. So he had a B there in Science 7. So what happened in that transition?" he questioned rhetorically.*

*Looking over the transcript with him, I noted aloud Jeremy's science marks from seventh grade onward, "B, D, F... F in his last [semester here]. That's interesting." Mr. Nuñez replied, "So you kind of see patterns. You know here, the pattern is sloping down.... So it gives you an entry point for your... you know, how to address this student. I mean, you can see his credits: 35, 20, 10, 7.5."*

*Mr. Nuñez continued to flip through the papers in the file. He pointed out the presence of Jeremy's standardized state test results from elementary school, as well as the amalgamation of English Language Development test scores dating all the way from Grade 2. Mr. Nuñez suggested that these two pieces of information combined could provide helpful background information on Jeremy's proficiency in English. He emphasized the School's goal of ensuring that all students are reclassified from ELLs to fluent English proficiency, the earlier in their school career the better. Flipping through the English Language Development test scores on record for Jeremy leading up to this academic year, he commented:*

*So by the time they get to us, you know, research shows that it's kinda' harder because they're older, so that there's more factors and barriers affecting them relative to when they're younger. So here you get a gauge for that, from like I said, [second grade], as a matter of fact, all the way to... to current.*

In his cursory review of Jeremy's grades and test scores, Mr. Nuñez began to get a general sense of Jeremy's academic standing and is able to pick up on some clues as to how Jeremy is progressing through his education. Although Jeremy's current grades would perhaps suggest low performance, Mr. Nuñez was able to see that, at least in the case of science, Jeremy had once excelled in that class in middle school. This led Mr. Nuñez to ponder insightful questions about Jeremy as an individual – what factors may have contributed to his downward turn in performance? Mr. Nuñez was also able to pick up on challenges Jeremy could have been

facing, particularly his long-standing ELL classification. Although not explicitly stated, it brought to mind the question, “With such prolonged intervention as an English learner, why hadn’t Jeremy yet been reclassified?” Mr. Nuñez showed me how he begins to assemble all of this information, forming an “entry point” for ways in which he might address this student in his own class.

While grades and standardized test scores contribute to a broad view of Jeremy’s academic standing, Mr. Nuñez turned through the file to discover some important qualitative data. “Um, so there’s a separate file, and this is the one I like because... it has teacher notes all the way from kinder to fifth grade. So, just to get some info HERE...” He began reading the documentation out loud, “‘Jeremy is a very happy student and has made progress in his social skills but needs to develop more self-control. He finds it difficult to settle down to the quiet routine of the classroom and to stay on task.’” Mr. Nuñez continued, “‘Fifth grade: Jeremy was in my classroom for 17 days. He needs to develop more self-control.’”

*Mr. Nuñez broke from the file to look at me. “So, without even reading the rest, you have that pattern of... you know, this is a fairly lively kid, if you want to put it in those terms. So then I would ask the new teacher, so how would... so let's say Jeremy has this habit. How would you address it? You know, what would you do?”*

Mr. Nuñez was able to show a progression from his own interpretation of Jeremy’s standard scores and grades – a student who was struggling academically – to one that was influenced by previous teachers’ experiences and exchanges with Jeremy. He was then privy to a documented pattern in Jeremy’s behavior, i.e., lack of self-control, that could contribute to Jeremy’s overall classroom performance.

Again, Mr. Nuñez did not apply this information as a predictive label on Jeremy. The assumption was not that Jeremy is intractably challenging and is anticipated to be disruptive in class. Rather, Mr. Nuñez attempted to incorporate these bits and pieces of information into a pedagogical strategy. If Jeremy does indeed have this behavioral habit, how might this be addressed in the classroom? How might a teacher proactively prepare to engage such a student?

Mr. Nuñez moved on to say that these pieces of documentation we had just reviewed are the most prominent elements of the cumulative file for his own analytical purposes. There are other supplementary items, however, that factored into his review, particularly as Jeremy's file begins in kindergarten and is fairly comprehensive. Mr. Nuñez flipped through immigration papers, indicating that Jeremy was born outside of the U.S.; early childhood development questionnaires measuring his "school readiness," which indicate that he started his LAUSD career at the earliest possible age; and, even samples of Jeremy's work in elementary school. Mr. Nuñez announced the title of each artifact:

*This one is "Why I love pizza." This one is "Hydroelectric Power," so that's science. "What were the dreams of Sally Ride and Louisa May Alcott?" Almost English, some sort of literature. "Comparing Plato and Aristotle." So it's a sample from various areas. This is like history, "Machu Picchu." You also get a sense of, not only their handwriting, but... their FRAME OF THOUGHT, you know? What are these kids thinking? Here's Dr. Seuss from Grade 6.*

The small portfolio of student work gave Mr. Nuñez further insight into Jeremy's approach to writing and content knowledge in his earliest years of schooling. Mr. Nuñez dug even a bit deeper, suggesting an even more personal connection with Jeremy, in the examination of his handwriting and "frame of thought" exhibited in composition.

Having moved through the entirety of Jeremy's file, Mr. Nuñez reflected on the steps we had just reviewed:

*I did this last year with a student teacher. And I gave him a quick rundown of, you know... what information that you want to take? You obviously want to get content, and you also want to get... definitely your content, any patterns in academics, [English Language Development] levels and any social/emotional skills.*

*Like for example, this Jeremy. I mean, you know we saw a few things like... self-control. Lack of self-control. So then, you think back to your classroom. Okay, is that still going on? If the problem persists or if it gets worse, then maybe a suggestion will be to refer to the counselor. You know, this kid has had issues for this long, I mean, is Mom aware?*

*Which she WILL be, you know, parents are aware. But have they... requested some sort of... help or second opinion, or medical opinion? ...Is it something that they can... does he have something? What's going on?"*

Amid all the seemingly disparate pieces of information contained in Jeremy's file, Mr. Nuñez attempted to focus on what was of highest priority to him as a teacher. He emphasized the importance of knowing the content to which Jeremy had been exposed, patterns in his past academic performance, Jeremy's English Language Development status, and any apparent socio-emotional skills or challenges. The repeated mention of Jeremy's issues with "self-control" by his previous elementary school teachers was data Mr. Nuñez kept in the back of his mind while observing Jeremy's current behavior. If this quality of character was persistent, Mr. Nuñez believed that this issue had been documented long enough to warrant intervention from the school counselor, Jeremy's parents, or even a medical professional. For Mr. Nuñez, the cume file serves as a track record from which explicit actions may be derived in order to support the individual student.

Not every file for every student is guaranteed to be as thorough as Jeremy's. If a student migrates into the District (from another district, state, or from another country), records may not be available prior to their admittance into LAUSD. In many cases, the transfer of the cume files

are not simultaneous with the transfer of the student. Teachers at Woodson College Prep, for example, mentioned that while cumulative files existed for the large majority of their inaugural student body, these files were not actually delivered to the school until four months after the start of the academic year. As a result, it was explained, Woodson began “with absolutely NO data.... We just opened the school without really knowing who our students were.”

Mr. Nuñez also emphasized that while files are available for most all of his students, the time it takes to intensively review records for every student precludes him from doing so for his classes, each of which are 30- to 40-students strong. Rather, he suggested targeting those students who “you think need the most help. Look at the students that have other issues apart from academics.” A student’s background, educational context, and history of intervention, Mr. Nuñez seemed to suggest, are most useful for those students apparently struggling with their overall academic performance.

Parsing out relevant student information from the cume files in this way is clearly dependent on each teacher’s discretionary diagnosis. The access and review of cumulative files is a completely voluntary undertaking by each teacher. Mr. Nuñez explained, for example, that some of his newer colleagues had heard of the files but had never seen them. As such, while the cumulative files are available (for the most part), integrating this information into classroom practice requires a certain degree of teacher capacity in determining whose files to access, how to access them, how to identify the most pertinent pieces of data as they relate to current student performance, and how to interpret data in the context of pedagogical strategy and out-of-class student support and intervention.

Also of note, Mr. Nuñez recognized that much of this information could be obtained via LAUSD’s learning management system, another primary source for him in reviewing student

transcripts, past state assessment results, English levels, and students' high school exit exam status. It may be that many teachers tended to rely on this online platform to obtain students' academic records rather than the physical files themselves (digital records were cited as a primary source of student data for the strong majority of case study teachers within this study). However, the online records are not always as comprehensive as the cumulative files, particularly with respect to samples of student work and teacher comments. And while low-technology, the physical files represent a mainstay of accessible student data impervious to the technical glitches that plagued the District's information system.

Examples taken from Belleworth School of Arts and Technology highlight how individual teachers made use of several data sources in understanding student performance. Mr. Nuñez focused on the value of understanding student background and community culture in building student-teacher rapport and, in turn, student engagement. Ms. Gavin underscored how knowledge of individual student strengths and weaknesses, as well as real-time information on student and classroom dynamics, are essential components of responsive instruction. Mr. Nuñez was also able to illustrate how he makes sense of standardized student data collected in individual cumulative files and the ways in which he incorporates these data into his own pedagogical approach. In all of these examples, external mandates to review or report student data, or even school-based administrative processes to do so, were largely absent. Rather, each teacher felt that these types of data routines were integral to their own successful practices and pursued them as a matter of course. Data indicative of individual student and classroom performance had been integral to the everyday practices of both teachers and would likely remain so irrespective of what changes in data use will occur at the school-level. Their direct

input into classroom instruction had, and would continue to be, essential to individual teachers as their pedagogy constantly interacts and reacts with different learners.

Interviews with additional faculty at Belleworth, as well as the two other case schools revealed that teachers independently determine what data they consider credible in evaluating student performance and in defining which instructional moves might affect improvements in student performance. From the calculations of grades, to the development of performance-based scoring rubrics, to the casual observation of students in their classroom environments (termed “kid watching” by one teacher at The Academy), there existed a natural propensity for teachers to identify, collect, interpret, and use some source of student data in their instruction.

The degree to which these processes are reliable, however, is debatable. Lacking formal, public systems of data disclosure suggests that teachers’ individual approaches to classroom data use can be prone to subjectivity. As an example, the interpretation of student data has been, for some teachers, a way of “predicting” low student performance based on assumptions or even demographic stereotypes. This stands in contrast to the ways in which teachers such as Mr. Nuñez and Ms. Gavin adapt their own instructional approaches to the content based on similar student data. In this sense, teachers’ interpretations of student behavior, progress, and potential are as individualized as the data collection methods themselves.

### **Woodson College Preparatory School: Credible Data Is Meaningful Data**

At Woodson College Preparatory School, data that are considered credible vary among stakeholders. For teachers, data that reflected students’ skills, capabilities, and progress were the most meaningful in terms of adapting classroom instruction to student needs. These included direct observations of students in the process of learning, as well as the collection of affective

data contextualizing student progress. Stepping back from student-specific data and taking a more collective view as to how well students were doing across Woodson's upper school was a larger challenge. In this case, the accumulation of student performance metrics, such as student grades, at the school level was not a type of data so readily embraced by teachers. This was due in part due to the difficult discussion of grading alignment required to produce an aggregation of student grades and the implications grading alignment was perceived to have on teacher autonomy.

### **Observational Data – Up Close and Personal**

At Woodson, the further away data were from representing individual student performance, the less valuable teachers found them. In part, this was linked to the perceived lack of validity aggregate student outcome data had in depicting a vibrant portrait of student progress. As an example, Mr. Macon discussed his review of Woodson's school report card annually issued by the District. For him, there was a distinct difference in the depth and quality of data produced for purposes of accountability and for purposes of instruction:

*We recently received our school accountability report card. And so that... in terms of accountability, we've definitely met, matched, and exceeded the District's... I want to say, requirements, maybe?*

*So attendance has been pretty good, expulsion rate is minimal compared to the District. In terms of whether the students feel safe here, same thing, it's a lot better than the District's average... So... on every point I would say that we're pretty good. We're pretty good.*

*But of course, that's not the only information that... the District can get from us to determine how well we're doing. The data that we're actually compiling right now, what we're doing with it, and how we are presenting it, and how we are disseminating it... THAT all matters to us. And how we communicate with our students, and what seminars mean, and what ADVISORY means, and what kinds*

*of things we do in advisory to help our students feel... self-directed, and active, and critical participants in society.*

*So we're trying to do a lot of the core competencies that we believe in, in our school. So those are a lot of things that AREN'T reflected by the District. You know it's more of the WHOLE CHILD rather than just numbers.*

Mr. Macon's feedback on Woodson's school report card evidences his review and understanding of its contents. He noted Woodson's standing in comparison with District averages and surmised that the School had met expectations of "competency" or "minimal standards." But having cleared these District-held "requirements," Mr. Macon turned to the data he found more relevant in determining how well Woodson was doing in its service to students. He pointed out that the School is actively compiling data that reflect its work in enabling students to become "self-directed, active, and critical participants in society" – core competencies also comprising Woodson's vision and mission. These were the pieces, he argued, that remained unacknowledged by the District. As a result, Mr. Macon believed that the District was missing out on a richer picture of the "whole child" and students' wider variety of core skills and capabilities. Without these data, the District's "accountability data" seemed to reflect "just numbers" rather than the real character of students as learners.

There was some value in the District's production of the school report card for Mr. Macon. Its credibility was held in the ability to show Woodson's progress against general indicators of performance and its comparison against District averages. This seemed to give Mr. Macon a sense of Woodson's relative efficacy. But to evaluate Woodson in the context of only these data feels uni-dimensional. Rather, Woodson had invested concerted effort in developing systematic data collection in and around character education, as well as to more thoroughly portray what learning through Woodson's courses and curriculum "mean" in context. In addition

to these unique data sources, Mr. Macon pointed out that how data are used and disseminated matter a great deal to Woodson's faculty. Teachers' participation in how these processes are carried out are not only an endorsement of the data, but also contribute to their credibility.

Furthermore, aggregate student outcome data are less useful to teachers as informants to instruction. When it comes to making decisions about curriculum and pedagogy, it is perhaps no surprise that teachers most value student performance data that come from direct observation of student activity. These observations, explains Ms. Lovell, are what gave her the opportunity to see moments of "growth." She provided a discreet example from the observation of a special education student the day before:

*Our geometry class, the kids work in collaborative groups. So [there's a] student who started reading on the second grade level... she probably can't multiply multi-digit numbers but... I overheard her talking to her group members about how to identify... like how to know if two sides of a triangle are similar--how you have to rotate it. And she was giving her General Ed peers... she was kind of explaining to them how you do that. So to me those are observational data, to ME that shows me she's making progress towards... mastering some... geometry standards. So that's good.*

*And the, also... a lot of observational data about kids' behaviors too. I know from making progress, just by seeing how a student is able to work well with others, or collaborate and cooperate, observational data is good.*

Here Ms. Lovell provided an example of how she captured students' academic and behavioral progress through direct observation. In this instance, incoming data presented to Ms. Lovell showed a student engaging her peers in an explanation of a mathematical concept. Not only did this offer Ms. Lovell the opportunity to see the student's grasp of the material in application, but also showed that the student was able to begin explaining it to peers through a problem solving exercise. Ms. Lovell tapped into the repository of information she knows about this student's current abilities (e.g., her reading level and math capabilities) and determined that

the student's exhibited classroom behaviors provided evidence of academic improvement. While Ms. Lovell drew on background data, as well as observation data, to evaluate her student's progress, she suggested that witnessing this stage of her student's development in understanding geometrical logic – the process of working through and wrestling with abstract concepts – was an assessment moment that could have only be captured through observation.

For Ms. Lovell, being able to observe learning as it occurs was a source of essential feedback for her instruction. Even in the context of reviewing students' written work, which she prioritizes as another valuable source of student data, Ms. Lovell explained that the most meaningful exercises are those that take place in class where she can see her students in the act of writing:

*I'm looking at students, I'm going around looking at their notebooks and I'm seeing they can't set up a ratio. Then... I can work with our General Ed teacher to stop the class and to kind of redirect them, or to point them into a different direction or to help them to see the... relationship of two shapes or something. So I can make some instructional moves to help push them.*

The opportunity to observe this work in progress allows Ms. Lovell to make immediate assessments of student need and redirect her instructional moves in calculated response. She contrasted this kind of instantaneous feedback loop to formal tests or student work that she collects and later grades. In her experience, the time lapse between the submission of student work and her provision of written feedback is too long for students to successfully apply her suggestions in an improved approach to the material. Ms. Lovell regarded tests as “informative,” but not “exciting” in the sense that they provide her with some valuable data, but not the type she can simply plug back into her instruction. She thought about this statement again, submitting, “I guess it's exciting if you know that [students have] learned the material. It's not that exciting if

they haven't learned it.” For Ms. Lovell, test results provide late notice of student achievement. If her students have performed well, she considers this exciting news. If test results are poor, however, Ms. Lovell feels deflated by a report that informs her, belatedly, that her instruction has not been as successful as she would have wanted.

### **Affective Data – More Than a Feeling**

In her consideration of what school-based data she finds personally valuable, Ms. Gilman also looked to observational data. The timeliness of observational data as feedback into her instruction, however, was less a focus than the ability to develop a nuanced understanding of individual students and to contextualize their performance and progress. Ms. Gilman began by acknowledging her value of more summary student performance outcomes, such as graduation rates, suspension rates, and reading levels. “Especially at the beginning,” she pointed out, “if you don’t really KNOW the kids that well... you can identify, okay, based on reading level, this person is going to need some serious extra support!” Standard student outcome measures can be useful, she suggested, in helping to identify student need, particularly if a teacher does not yet know her students very well. But for Ms. Gilman, the point of teaching *is* to know her students well – well enough to identify and evaluate what kind of progress they have made as learners. In this sense, she places a great deal of credibility on observable in-class performance.

She provided as an example one of her students, Cecilia, who was new to Woodson at the beginning of the year and who led a very “quiet and introspective life.” Ms. Gilman recounted, “I wouldn’t really know any of her ideas except if she wrote them down and turned something in.” In addition to Cecilia’s introversion, Ms. Gilman distinctly recalled that, despite having immediate family from Mexico, Cecilia was unable to locate Mexico on a world map. By the end

of the year, Ms. Gilman had watched Cecilia grow as a major contributor to class discussion and debate. “She’s not going to talk A LOT, but she is DEFINITELY going to talk. When she does the whole room kind of gets quiet because they know she’s going to say something SUPER DEEP.” Ms. Gilman amused herself in this assessment, noting, “That’s hilarious. Like that’s data: the room gets quiet when she begins to talk.”

Although Ms. Gilman found some humor in what data she identified as credible in estimating Cecilia’s personal growth, she is consistent in her approach to measurement. Her observations of Cecilia’s classroom participation not only presented evidence of increased participation, but what Ms. Gilman saw as a “growth in thinking” and a transition from being someone who was “sort of quiet and *not known*” to someone who was obviously attractive to other students as a group member who will help them to do well in class. Further still, Cecilia’s engagement in classroom geography games, such as one in which students were prompted to identify different countries on a map, indicated that Cecilia was practicing “all the time” at home. Ms. Gilman enthusiastically remarked, “I’m a geography teacher. I struggle to say all of the countries in Africa and where they are... But [Cecilia] has been practicing Africa for, like, *weeks*... And she got 90% in African geography! This woman who didn’t know where Mexico was! It feels so *significant*, you know?” Through her observations of Cecilia’s classroom performance, Ms. Gilman had pieced together a rich picture of Cecilia as a contributing member to her classroom community, a thinker who was held in great esteem by her colleagues, and a diligent, hard-working student. An assessment of Cecilia’s geographical knowledge might well have been conducted via test form. But what Ms. Gilman pointed out was that Cecilia’s content knowledge was, importantly, characterized by “the way she [was] seen.” Ms. Gilman continued, “I think I would even say, some of the way that she sees herself has really GROWN and

*developed.*” In this way, Ms. Gilman’s understanding of Cecilia’s growth in character, disposition, and level of engagement with classroom content could have only been ascertained through classroom observation.

### **Enhancing Intuition**

Ms. Gilman’s comment regarding Cecilia’s improved geography quiz scores “feeling significant,” however, lends itself to some scrutiny. It seems strange that, even though Ms. Gilman was able to empirically measure Cecilia’s content knowledge, she continued to base her determination of Cecilia’s academic growth on “feeling.” In part, Ms. Gilman may have been referring to her own extended excitement over Cecilia’s progress. But her comment raised the question of whether direct observation of student work and behavior as a genuine data source is undermined by natural human inaccuracy and subjectivity. Ms. Lovell addressed this issue in considering her own approach to observational data. She noted that an area in which she would like to improve is being more “systematic” about her observational data, “having a better lens and really being more... cognizant of what I’m actually going to observe.” She explained why this sharpening of her focus is so important:

*I think I need to be better at KNOWING what I'm looking for in students. Because sometimes what happens is one student, maybe two or three students, are doing well... observationally. Like, they're engaged, they're talking, and that a lot of kids AREN'T. But because those three kids are, it shapes my experience as a teacher. I feel like it's GOING well because of having a class discussion with only three people, but it feels like it's the whole class as a teacher. Because the other kids are looking and they look kind of like they're listening.*

*So I guess, like, taking better notes, or being able to better know what I'm looking for.... Really, who am I calling on? The four kids that are always talking? Are other kids taking notes, writing, listening... can they contribute? Things like that. Like being more... systematic. Not systematic... being more... PURPOSEFUL.*

Ms. Lovell made a critical distinction between systematic data collection and purposeful data collection. It is not just that she saw the need to regularly go through the motions of recording and reviewing what she observes in her classroom. Rather, Ms. Lovell underscored the importance of targeting her observational data scope in response to specific questions of practice. Although she might have gotten the sense that her class was participating in classroom discussion for example, it might very well be she was only actively engaging a handful of students. The purposeful collection of observational data might have helped her assess whether non-verbal students were indeed participating through writing or listening. Reviewing these data might help Ms. Lovell think through ways of increasing student contributions to class. Without this empirical data collection, however, it might be easier to assume that the whole class is participating based on the involvement of just three students who naturally “shape her experience as a teacher.”

From Ms. Lovell’s perspective, observational data are essential to understanding student ability and progress, and there are some ways to improve the collection of these data in critically examining her own instructional moves. However, she also expressed some frustration with the limited credibility assigned to affective data that are purposefully collected. As an example, Ms. Lovell cited her work on a survey designed to gauge students’ engagement in, and value of, a seminar program geared toward career preparation. In general, she didn’t feel as if the survey returned “great information,” much of which was in the style of self-ratings on a sliding scale. Personally, she felt that the best questions were students’ free responses:

*Because the kids would then say, like, “This space is really out of the box. These are the things that I’m learning.” But it’s so hard to measure terms of like NUMBER. You know what I mean? And a lot of data you want to see some sort of increase of SLOPE. I mean, it was really hard to measure....*

When asked her perception of why quantitative results, and in particular the “slope” was so valued and necessary, Ms. Lovell replied:

*I don't know, because that's what the District people like. It's really annoying.... That's what they always look for.*

When asked if this was requested from her all the time, Ms. Lovell responded:

*No, but that's what they always look for.... You know, [your] school is amazing because there's an upward slope of the line.*

Although students' open feedback on the seminar program survey was most valuable to Ms. Lovell in terms of gauging the program's success, she felt that these type of data were habitually not viewed as credible for those evaluating school performance. In her experience, some quantitative measure of pre-test to post-test improvement is the only acceptable evidence of growth. Ms. Lovell stated the issue quite clearly: many educational outcomes are hard to measure. But for Ms. Lovell, as well as many other study participants, until there a definitive way to measure how well students and schools are doing across a variety of outcomes comes to fruition, the educational community must acknowledge the variety of data sources contributing to this complex understanding.

### **Grades Ain't Nothin' But a Number**

When asked what accountability data were requested of Woodson by the District, the principal, Ms. Figueroa, listed student grades as one of the big categories. Grades, she explained, are the basis for understanding whether and how many students are passing classes and moving through courses required for graduation, as well as graduation and college acceptance rates. Although likely one of the oldest metrics of student performance in educational history, Ms. Figueroa understands grades as a complicated measure and the product of a compound

construction of meaning. She explained, “If you focus on instruction, you have to focus on what students are really learning, and you have to focus on what this grading really means.” But because of this relationship between grades and what they represent in the context of instruction and student learning, Ms. Figueroa had also found resistance among Woodson faculty in using grades to assess school performance. When it comes to reviewing and discussing student grades as a school, Ms. Figueroa remarked, “That's a BIG hot button. Nobody wants to talk about how they grade or what matters to them.”

Dr. Baher, Woodson College Prep’s residence researcher, shared Ms. Figueroa’s frustration with faculty’s refusal to examine student grades as a metric of school performance:

*Where I think they could be more mindful and more critical is... the course failure rates. That's data that continues to trouble me. Because way too many kids fail classes. And that's not unique to our school, but it's something that... I know WORRIES teachers. And I know teachers don't fail kids lightheartedly, that's not what I'm suggesting, but it's hard to have a conversation about that.*

*So more than once, I've tried to develop, I've had [Professional Development] conversations about grading. What does it take to pass your class? What does it mean to get a C? How do you... give out grades? How much does homework count? And those are VERY hard conversations to have because teachers have just an enormous feeling of... like, that's MY GROUND. Right? And you can give my kids a test, but that, those course grades, they're mine.*

Dr. Baher pointed out a perplexing tension: while Woodson’s teachers obviously care about the success of their own students, they show difficulty responding to Woodson’s high student failure rates. Although no teacher fails his/her students “lightheartedly,” Dr. Baher has met great challenge in constructively discussing how this might be resolved. This is because, she suggested, the question of grading inherently questions teachers’ instructional approaches. Asking teachers to negotiate minimal performance outcomes for their classes, how they prioritize various demonstrations of student ability, and how these align with teachers across classrooms

and departments is viewed as a serious impediment on teacher autonomy. It is one thing, Dr. Baher suggested, to assess student knowledge through standardized exams, but it is quite another to standardize a grading structure. “It’s a very interesting conversation,” she continued, “because I think teachers feel like grades are so tied to their professional credibility, and their judgment, and their autonomy.” From this perspective, student grades are only regarded by teachers as credible if they are left intact – as each teacher has intended them. At the school level, however, student grades lose their credibility as accountability measures because, “they don’t mean the same thing school-to-school, class-to-class.”

Teacher sovereignty associated with grading practices is in some ways related to a sense of authority in the classroom (“this is my ground”), as well as a sense of flexibility teachers feel they need to address unique classroom needs. In a discussion of his own department’s grading practices, Mr. Macon explained how he and his colleagues have discussed the meaning of “basic standards” in science, such that a student who earns a 70% might be considered proficient. But, how each teacher composes that 70% is left to individual determination. “For example,” Mr. Macon offered that his “quizzes aren’t weighed as much as my exams. My quizzes are only 5% of [students’] overall grade. So, in that case, my students could get away with not doing well on those quizzes, but doing overall very well in the class.” Despite the Science Department’s common understanding of what basic standards students should obtain, how grades are assigned is still left to the discretion of each teacher.

There are some places where Woodson’s teachers were slowly beginning to cooperatively structure approaches to grading. The Math Department, explained Dr. Baher, had been a front runner in its discussion of mastery-based grading and had open conversations about common expectations for grading. But for the large majority of teachers and departments, this

was not the case. Interestingly, some teacher participants seemed less enthusiastic about grading data, not because of their concerns for teacher autonomy, but because student grades were considered less relevant to their instructional practice. Even with some discussion around student proficiency in the Science Department, Mr. Macon interjected, “that doesn't help me though, with the data I'm getting, in terms of HOW I can help.” For Mr. Macon, grades may help to identify a student’s overall level of proficiency, but they do nothing to inform how his instruction might actually be modified to encourage improved student performance. Similarly, Ms. Lovell questioned how informative student grades are for her own approaches to teaching and learning:

*Final class grades, I use that to tell me... whether or not the students are... GENERALLY succeeding in the school system. Because I guess, still to me the grades still don't really reflect what they know, but it more reflects their like SCHOOL skills.*

*Like can they complete an assignment, can they comply with, you know, teacher requests, can they like organize themselves enough to finish something? Do they have the smarts to ask people for help or get the resources they need to figure it out? To me that's what a grade represents.*

In a strange, self-propagating cycle, these attitudes toward student grades seem to both result from an absence of teachers’ cohesive understanding around grading and serve as a source of disinterest in using grades as a metric of school performance. That is, teachers find less value in student grades because they are not regarded as an accurate reflection of student knowledge or because they do not offer teachers insight into potential instructional improvements. And because teachers do not find practical value in student grades, there seems to be little interest in having further dialogue about how to link grading practices with expected learning outcomes, or to ensure that grades have transferrable meaning across subjects and grades.

Several sources of data maintain credibility at Woodson. Data collected directly by teachers in the course of classroom activity, as well as more aggregate-level performance outcome data, were all regarded as credible sources of information. However, teacher participants distinguished what weight they allocate to each data source, giving clear priority to data that are most useful in affecting their engagement with students. For Mr. Macon, understanding the context of what students are expected to learn (not just what they have learned, but for what purpose and through what processes), data sources used to evidence school and student performance must extend beyond standard accountability measures published by the District. For Ms. Lovell and Ms. Gilman, the direct observation of student learning – as instantaneous feedback to instruction and as a portrayal of student progress deeply embedded in context – is most valuable. Because of their value for these data, some participant teachers at Woodson looked to reinforce their credibility through systematic collection. As heard from Mr. Macon, this can take the form of involving teachers in data routines built around school-specific indicators of performance. Observational data might be collected in more purposeful ways, considered Ms. Lovell, who thought a more structured approach to classroom data collection might reveal aspects of student participation she had not yet detected.

On the other hand, experiences with grading data at Woodson presented a much more complicated picture of data credibility. In this instance, teachers' antagonistic relationship with student grades fueled a reticence among faculty to respond to high student failure rates. Per Woodson's principal and resident researcher, teachers only found grade data credible when they were in control of how they are issued. At the same time, teachers did not deem school-level grade data as credible because the grading practices of other teachers did not mirror their own. Additionally, discussions with teacher participants suggest that Woodson's teachers did not

necessarily find grades useful in informing instruction or in accurately portraying student content knowledge. But it is perhaps because of this general disinterest in grades, that teachers were also less responsive to conversations about enacting changes to their grading practices.

Various stances within Woodson toward data credibility indicated a deep-seated divide between classroom-level and school-level data. School-level data were used to assess Woodson's general performance, both in terms of District accountability and formative school improvement. Underlying this introspective view to schools was the theory that the use of data to identify school shortcomings and successes would lead to targeted conversations about how to improve school performance. Such improvements naturally imply revisions to instruction. But despite this chain of inference, teachers did not necessarily see themselves, their work, or their students in aggregate data. Rather, these data were seen as abstractions of classroom practice and student achievement removed from the actual process of teaching and learning.

### **Cross-Case Insights**

This chapter has discussed the many different types of data school stakeholders identify as “credible.” Examples from The Academy show that, even when formal data infrastructure does not yet exist, faculty have been able to identify data sources reflective of their mission, vision, and school culture. As such, The Academy's school culture was reciprocally defined by the data upon which it had endowed credibility. These data include collaborative teacher assessments of students' academic and behavioral performance, as well as teacher performance. Both systems of review focus on the collection of individualized, contextualized, and nuanced data to develop a full-bodied picture of progress and growth.

Examples from Belleworth highlight types of credible data participant teachers rely on to inform their pedagogical approaches. These include an understanding of students' community and culture, as well as students' personal strengths, weaknesses, and orientations to learning. One teacher walked through how he filters through students' cumulative files to make sense of routinely collected student data, such as ELL status, grade reports, teacher-developed progress reports, and examples of student work. Whether these teachers are reviewing standard student outcomes or gathering data from students through classroom interactions, in all of these instances, credible data were those specific to individual students and which provided insight into students as distinctive learners.

Participant teachers at Woodson attributed some value to school-level data in identifying broad patterns of student performance and areas for improvement. But aggregate measures of accountability were considered less useful than data closely examining the processes of teaching and learning underlying student achievement. This was attributed to teachers' interest in exactly how to affect student achievement, the need to make real-time changes in instruction, and the desire to understand student progress in the context of individual students' learning experiences. While specific data collected directly by teachers were viewed as essential in altering instruction, its prioritization had also been seen to undermine teachers' constructive reflection on school-level data. For example, teachers seemed to struggle with addressing overall student failure rates through collective discussions about their own grading practices and standards of learning.

Across all three cases, the credibility of data is reliant on what meaning and what value users confer upon data. Interestingly, aside from one teacher participant at The Academy who didn't believe standardized testing "is beneficial for anyone or really shows anything that's true," participants did not actively discredit as wholly invalid any data used by schools and the District

to assess student and school performance. It was more common that teacher participants felt the data they most valued were largely unacknowledged at the District-level. Many participants expressed frustration at the failure of accountability data to capture the nuanced experience of student progress and performance. This is probably due in part to practical limitations in compiling aggregate-level data used to assess school performance across the District. It is extremely difficult to construct common indicators of school effectiveness that are transferrable across schools and simultaneously sensitive to individual school contexts. Because of this, however, teachers experienced a lack of connection between accountability data and their own practice, despite the implications accountability data had on instruction. As such, some data were used for instruction, other data for accountability, and credibility appears to be strongly linked to their disparate uses.

## CHAPTER 7

# CULTURES OF DATA USE

### Introduction

The discussion of data credibility has focused on the various types of data that school stakeholders identify as measures of student, teacher, and school performance that are accurate, meaningful, and trustworthy. With so many different types of data available to schools – formal and informal, systematically collected and unsystematically collected, quantitative and qualitative, aggregated and disaggregated – eventually data are selected to inform decisions around instructional practice, student progress, and school improvement.

Importantly, the consideration of what data are *credible* is not without some reflection on what data are considered *useful* by different groups of stakeholders. Data credibility is in many ways dependent upon the perceived or anticipated application of data within a school's context. The intended use of data can, therefore, be an influencing factor on whether data are regarded as valid or reliable. For example, in Chapter 6, we examined how teachers at Woodson College Prep considered student grades credible data in identifying general areas of student need, but less credible in the evaluation of Woodson's overall effectiveness. How data are regarded by school stakeholders is, therefore, intertwined with a discussion about the many purposes for which data are used.

Looming even larger is the question of how data are utilized within processes of decision-making – the substance of this chapter. That is, how are data integrated into schools' conversations about teaching and learning, if at all? What are the ways in which these data are seen to influence stakeholder perspectives? How, if at all, are data used to substantiate the

outcomes of decisions? Chapters 4 and 5 have explored how data use is dependent on basic data structures and systems in place within a school, the research questions a school is trying to answer, who is responsible for decision-making, and how decision-making is pursued. This chapter presents examples of how data are or are not applied within processes of school decision-making, and is presented in three parts. Part I looks closely at factors that support and impede data use in strategic development and instructional planning across two school sites. Part II presents an in-depth analysis of the use of student assessment data to inform curriculum and instruction across several departments at Woodson. Part III specifically explores examples of how the use of performance data in all three school sites introduces tension between teachers' sense of mutual accountability and personal autonomy.

## **Part I: Data Use in Strategic and Instructional Planning**

In the attempt to understand how data are integrated into decisions involving school program development and implementation, as well as instructional planning at the classroom-level, examples are provided from two of the three school cases. Teachers within The Academy make use of self-collected data to inform their own classroom activities, and the School is making early preparations for its self-study required for accreditation. However, The Academy is still establishing and routinizing systems of data collection. The focus of this section, therefore, is on Belleworth School of Arts and Technology and Woodson College Preparatory. The case of Belleworth provides an example of the ways in which student outcome data are used to garner support for student interventions. Leadership within Belleworth soon discovered that the use of data requires faculty to make a personal connection with the data as a way of seeing their students in “the numbers.” The case of Woodson focuses on its implementation of an

“Improvement Science” initiative wherein teachers collect, analyze, and interpret classroom data for use in instructional improvement. In addition to discussing some of the benefits and challenges of the formalized use of data in the classroom, the experiences of Woodson’s teachers highlight some of the challenges associated with a focus on measurement. Faculty members also bring to light important considerations in the conversation around data use as an integral component of instruction.

### **Belleworth School of Arts and Technology: Using Data to Guide Program Development and Strategic Planning**

In her new position as principal, Ms. Heredia was excited about Belleworth’s unique opportunity as a pilot school to tailor its structures and systems around student need through the exercise of its autonomies. She considered herself an advocate for the use of data in guiding the development of school programming and student support interventions and looked to data as a foundation of evidence from which Belleworth could construct its defining approach to instruction. The “big picture” question she believed Belleworth needed to answer was: “What are you doing differently?” Strategizing in this way presented a new approach to administration at Belleworth, requiring a shift in perspective for many teachers, and in particular, Belleworth’s ILT:

*Because we’re a pilot school and we have autonomies, what I’ve been pushing [the ILT] on is how are we using those autonomies? Like, what is our evidence to show how it is we’re using our freedoms? And we aren’t.... Besides being kind of like a smaller version of a comprehensive [school], what we’re doing is what’s being done at a comprehensive, you know? So... part of the work this semester has also been, how are we going to use this data to then think of our autonomies to fix it? Like, how is this data going to lead us in... how effectively we want to use our autonomies? And that’s what we’re... trying to think, like what are we going to do different? And not just different to be different, but different to address this need.*

Ms. Heredia emphasized that, in order to distinguish itself from a traditional, comprehensive high school, Belleworth needed to capitalize on its freedom to self-govern. She emphasized that the goal was not just to make Belleworth distinctive for the sake of standing out from the crowd, but rather to strategize creative approaches to addressing student need. Identifying those needs, she argued, is reliant upon the analysis of Belleworth's student performance data. In this way, the data "lead" the school into thinking about how it might most effectively make use of its autonomies. Data should also be used, she argued, to evidence Belleworth's successful exercise of its autonomies.

### **Using Data to Inform Student Supports and Interventions**

One major initiative Ms. Heredia looked forward to implementing in the upcoming year was the revision of Belleworth's bell schedule to accommodate an additional period of instruction where students would participate in either intervention or enrichment activities. Reviewing the percentage of students passing all of their classes, Ms. Heredia noted that while there were slight increases in this rate over last year, a substantial proportion of students were still not passing all of their classes. As a way of attending to this serious issue, Ms. Heredia needed to develop a multi-pronged approach: 1) work with her ILT to brainstorm how Belleworth might better support failing students; 2) work with the entire faculty in understanding student failure rates at Belleworth; and, 3) mobilize the school to implement newly-devised student interventions.

Putting heads together with her ILT, Ms. Heredia and several other teachers looked at students' grades from the 5-week, 10-week, and 15-week grading periods. Collectively, they decided to use a Response to Intervention (RTI) approach to categorize students into three

performance “tiers.”<sup>6</sup> “Tier 1” students were those passing all of their classes, “Tier 2” consisted of students failing one or two classes, and “Tier 3” students were those failing four or five of their classes. An RTI Committee was formed (ultimately mirroring the membership of the ILT) and charged with, as one committee member explains, “looking at student data, capturing student data, and sharing it with the staff.” Varying degrees of intervention were discussed for each student tier.

One of the first interventions trialed with Tier 2 and 3 students, for example, was mandatory afterschool tutoring. Teachers were asked to select five students who were failing their class and tutor them for an additional hour each week for a five-week grading period. Throughout this period, the RTI Committee tracked student grades and distributed a teacher survey requesting feedback on the tutoring program including whether teachers felt student participants were improving in their classroom behavior, overall engagement, and/or academic standing. The tutoring sessions received mixed faculty reviews. While some teachers believed the program contributed to improved classroom behavior, other teachers were found not to hold regular tutoring sessions as expected. The inconsistent implementation of the tutoring program led to its suspension, but the data collected from this initiative fed into the development of a redesigned bell schedule.

---

<sup>6</sup> The Response to Intervention program is an official program of the RTI Action Network and the National Center for Learning Disabilities. It is described as “a multi-tier approach to the early identification and support of students with learning and behavior needs” where “struggling learners are provided with interventions at increasing levels of intensity to accelerate their rate of learning.” RTI relies upon the analysis of student performance data, and is “designed for use when making decisions in both general education and special education, creating a well-integrated system of instruction and intervention guided by child outcome data” (National Center for Learning Disabilities, n.d.). Belleworth’s interest and involvement in the program drew from materials produced by the RTI Action Network, although the ILT had not yet attended any formal professional development events hosted by the program. As such, Belleworth’s implementation of RTI was an independently driven initiative.

In this new master schedule, the day was expanded from six periods to seven and, based on their individual learning needs, students would be placed into either an intervention program or an enrichment program during the additional period. Defining a systematic placement process in accordance with each student's needs required an intensive process of student review and involved several additional sources of student performance data. As explained by Ms. Gavin, also an RTI member, Belleworth's English teachers were first given the full list of students, from which they identified those needing "intervention" in their subject area. The list was then given to Belleworth's Math teachers who underwent a similar process. Additional "core content" teachers identified students requiring intervention in yet a third round of review. This iterative process of student placement was intentionally designed to ensure that it would draw upon a strong base of data, in this case, teachers' intimate knowledge of individual student ability and performance. Ms. Gavin detailed:

*We had to figure out, should we do it through Advisory? Like should I as an advisory teacher go through my kids and say yes or no? And then we decided that that's not personal enough. It needs to come from the content teacher FIRST. Because the content teacher KNOWS whether they need intervention in that content or not....*

*We needed to clarify. We need to figure out what they mean by "they're not doing well." And that's one thing that we talked about yesterday: is it a behavior problem, or is it an academic problem? So that's another thing to consider.*

*So AFTER they've done that, eventually it will come back to that homeroom teacher that should KNOW them. And they say, "So-and-so has been identified as needing intervention in this, this, and this, in this course. Do you agree or do you not agree?" So it's gonna' eventually come back to... so basically it's kinda' like, just because YOU think they need intervention, more people's eyes are going to look at that kid to either agree or disagree, and then you're going to have this discussion with your grade level team to decide. So before next year, all of the ninth grade teachers will sit down together and say, "You know what? I think that kid DOES need intervention."*

Ms. Gavin here explains a highly-detailed process by which each student is individually reviewed by his or her teachers in the determination of whether they are placed on the side of intervention or enrichment. Interestingly, while the RTI Committee had been focused on the use of student performance data to substantiate program development, she underscored the importance of consulting teachers who “know” their students most closely in order to inform this decision. This was not an algorithm based upon grades – who has passed and who has failed – but a close discussion of where a student needs to be along a continuum of “doing well” to “not doing well” in terms of both behavioral and academic performance. In this way, Ms. Gavin acknowledged the importance of relying on teachers’ professional judgment. She emphasized a collective stance toward appropriate student categorization, wherein teachers serve as checks and balances for one another. In this particular exercise, data is drawn from both teachers’ personal systems of student assessment and evaluation and the ways these are collectively interpreted and negotiated among Belleworth faculty.

This is not to say that regularly documented and tracked data, such as student grades, have been sidelined. Much to the contrary; eliciting staff buy-in to the notion of a new bell schedule (representing the third bell schedule revision for Belleworth in four years) necessitated a deeper faculty understanding of student failure rates. As Ms. Heredia noted, data “set the stage” for the proposal of the new bell schedule:

*We have, I think, several good ideas. It’s just a matter of seeing how they work on a master schedule, seeing what that bell schedule would look like, and then getting teachers to approve it. So this is why this work is important right now, because if you show teachers this information, how could they not, you know?*

*So we know this, we know this, we know this. We have data about ALL of that. And so you’re not going to vote for this WHY? “Well, ‘cause we don’t want to have to stay here longer or because we don’t have to have that extra prep.”*

*So this is the student need, and this is the teacher need. So you're not going to vote for this WHY? You know what I mean? You got to set the stage with your data as to why you need to do what you need to do.*

Ms. Heredia describes using student performance data (and high student rates of failure in particular) almost defensively in her advocacy for a new bell schedule. Her comments above allude to the pushback received from some teachers who are disapproving of the new schedule because of its addition of another class to their current workload. But for Ms. Heredia, low student pass rates present a clear picture of the path to be followed. To ignore such salient evidence of student need is morally indefensible, even if it means discounting teacher preference. In instances like these, Ms. Heredia believes that the data make the strongest argument for action and are imperative in framing deliberation and discussion around Belleworth's instructional strategy.

### **Forging Personal Connections With Data – A Prerequisite of Data Use**

In order for these data to carry any weight with Belleworth's faculty, however, there is also the need for teachers to develop a sense of internal accountability to the facts and figures. As Ms. Heredia put it:

*Every time [students] fail a class, they're going to be demoted; they're going to stay behind a year. And I was trying to get teachers to connect. These numbers matter. Because sometimes they think when you're data driven that you're not thinking about the whole child. But I'm like, this IS a child who is not going to graduate.*

Here, Ms. Heredia addresses a concern among some faculty that student outcome data do not comprehensively encapsulate student potential and aptitude. In the case of failure rates, however, Ms. Heredia considers such an argument a poor interpretation of the one-to-one correlation between failing a class and a student's subsequent demotion. The goal for Ms.

Heredia is not simply to think of data as numbers bearing punitive weight, hammering faculty with disheartening student statistics. Rather, she is hoping to engender a personal connection with figures like failure rates. These are not just numbers, but numbers that “matter” in their representation of individual student success.

In attempt to establish a “connection” between student outcome data and teachers’ approach to instruction, Ms. Heredia worked with the ILT to develop an exercise wherein teachers were given time to reflect on student failure rates. Stacks of student files were prepared for each teacher of those students who were failing only their class. During a staff meeting, teachers were given their stack as a physical representation of the students they were failing. She explained teachers’ reaction to the activity:

*So these kids are passing everything else but you, you know? And so, ONE, it was eye opening because they were like, what? This kid’s passing everything? Because the teachers don’t know sometimes what grades kids are getting in other classes. Teachers don’t know. It’s like in isolation.*

*So a teacher will assume they’re failing my class, they just suck as a student. When it’s like, wait, they’re getting straight A’s? Like, what? What am I doing wrong? Or I had teachers that had stacks this much [shows width with hands] of kids failing and other teachers that had no kids failing, and it’s like wait, what? Why do I have so many? You don’t have [looking around]... oh. [Short laugh] You know what I mean? I need to figure something out. Like, what’s not working in my classroom?*

*You know the process is simple, but what it does is a lot. Because it at least gets teachers thinking things need to be different.*

Ms. Heredia described this activity as an eye-opening experience for her faculty, particularly for those teachers who had relatively higher proportions of students failing their classes. By being given a physical representation of those numbers, i.e., the stacks of student files, teachers were able to glance around the room and immediately compare the proportion of

their failure rates with those of their colleagues. Ms. Heredia saw that this was a first-time opportunity for faculty to reflect on how their students might be fairing in other classes, and in turn, reconsider their own grading practices. Teachers were forced to ask the question, “If it is the case that so many students are failing only my class and no others, what does this say about my instruction?” As such, the challenges of student failure rates became less a problem of abstract numbers and more an issue requiring personal reflection and involvement.

Ms. Salçada, a member of the ILT, underscored the importance of this exercise for re-orienting teachers’ view towards failure rates:

*So I pulled the reports, so these are the kids who were passing all of their classes except for your class. So it's not an academic challenge, it's probably not socio-emotional, it might just be something that you can address within your classroom. Because it's not like an academic challenge that's impacting all of the classes. So it's something academic that you can probably address in the classroom. There's obviously some potential with this child. So how do we support this child to be able to be successful and pass all of their classes?*

In this excerpt, Ms. Salçada sees teachers’ confrontation with their “stack” as a way of stripping away general assumptions about student performance. If the student had difficulty in an academic setting, or is lacking a foundational academic skill set, for example, the expectation would be that they would be failing several classes. Larger patterns of failure across classes would also likely be present if the student had substantial socio-emotional needs. However, if a student is failing only one class, there must be something that can be done within that class to support their improvement. Even in the context of just one individual student, this line of thinking compelled teachers to reflect on the interaction between their own instructional approach and that of a particular learner. As Ms. Salçada later put it, there must be something in a teacher’s “grading or classroom practices” that accounts for the “discrepancy” in wide

variations of student performance across classes – that there are, in fact, “a lot of questions that we need to look into.”

Ms. Salceda later explained that, as she continues to track students’ academic standing, if she sees that teachers are still failing large numbers of students over the course of the semester, it will likely warrant a conversation on a one-to-one basis. Within these meetings, she anticipated asking teachers more directly, “How are you supporting your kids to be able to meet your expectations? Why is there such a big gap that kids are passing all of their classes except for yours?” Her plan was to walk through teachers’ class-based failure rates as way of holding them accountable to their own students.

While some teachers may be reticent to use student grades as a metric of performance (such as some teachers from Woodson who perceived grade calibration as a subject of non-interest), Ms. Salceda sees these data as an important reflection of teachers’ collective work within Belleworth and a valuable data source. Student grades are something that Belleworth faculty find credible because of their active role in defining them:

*So this is, like, our data. We created this data. So now, how do we make it better? How do we change things? So it's been a lot of days, a lot of work. It's not like an external organization coming in to tell us, “Well, you guys did great at this, but you suck at this.” It's something that was created from within.*

Rather than rely on “outsider” interpretations of how well the school is performing, Ms. Salceda believes that the review of teachers’ internally-developed grades grant them strength in validity and, in turn, are worthy of use in decision-making. Because grades are co-created by Belleworth’s students and teachers, they are robust under scrutiny and can be used to substantiate questions of monitored improvement. Teachers’ active role in constructing the data thus implies

their ownership of the results and their propensity to use such data in gauging school performance.

The development of support programming for Belleworth's students was not a new activity. Targeting interventions and leveraging pilot school autonomies to more effectively address student need using student achievement data, however, was a recent introduction into Belleworth's process of strategic planning. As an example, this chapter highlights how student grades and low class pass rates have been used to substantiate a new master schedule and the addition of a class period reserved for student enrichment or intervention activities. These data have been used not only to highlight achievement trends and compel teacher support for a new master schedule, but also to evidence prominent student failure and to defend the moral argument for increased teacher effort implied with the additional period. Additionally, teachers' personal assessments of student progress are also data taken into consideration in the sorting of students between enrichment and intervention programs. Belleworth's process of determining student placement would involve a systematic review of individual and collaborative teacher evaluations of student achievement.

The infusion of data into the process of decision-making within Belleworth was not immediately accepted by faculty. Connections between student achievement data and teachers' classroom practices needed to be forged as a way of personalizing the data prior to their analysis and interpretation. Rather than view class pass rates as a numeric abstraction of student performance, for example, Belleworth's ILT and principal, Ms. Heredia, worked to show faculty how "the numbers" reflected individual students. By comparing how many students were failing only their class, teachers were given the opportunity to consider how their own grading practices, and potentially their instructional practices, might contrast with those of their colleagues in

supporting student success. This exercise was seen to be successful in part because of the regard Belleworth faculty have for their grading data. Because of teachers' central participation in grade creation, these are data that Belleworth's faculty both understand and endorse in their use for instructional planning.

### **Woodson College Preparatory School: Using Data to Guide Classroom Instruction**

While Belleworth was just beginning to introduce the analysis of school-based data into its strategic planning processes, this was an ongoing objective for Woodson College Preparatory School. Woodson's concerted focus on creating meaningful data for the purpose of guiding instruction, as well as the many metrics it annually publishes as evidence of its progress and achievements, are indications of an administration and faculty adept at using data in decision-making processes. Woodson is not unlike Belleworth, however, in its continuous endeavor to mobilize teachers around standardized student achievement data. Like Belleworth, the success of Woodson's data-focused initiatives was predicated on an organizational value for data in understanding student performance and teacher practice. It was also dependent on an adaptation to data conventions by individual faculty members and their own understanding of how, and for what purposes, data were collected. Consequently, Woodson's ability to elicit individual teachers' buy-in and meaningful engagement in self-developed student assessments (or common assessments) and improvement science initiatives had been a gradual process.

### **The Science of Improvement**

Woodson had been working for several years to implement a method of data collection, interpretation, and use as represented by the field of improvement science. Within each department, teachers were expected to engage in rapid, iterative cycles of evaluation whereby

teachers independently would: 1) *Plan* – i.e., identify an area of student growth for their classes, figure out a “root cause” of that growth, and identify a “change idea” to cultivate growth; 2) *Do* – i.e., collect quick, formative data (i.e., “run data”) before and after implementing their identified “change idea;” 3) *Study* – i.e., analyze “run data” in the determination of whether the “change idea” made a difference; and, 4) *Act* – implement new or modified “change ideas” alongside the collection of additional “run data.”

Termed Plan-Do-Study-Act, or PDSA, by Woodson’s faculty, these short, iterative rounds of evaluation provided a framework for department-based strategic planning and professional learning. Yearly PDSA data were even meant to contribute to teachers’ portfolios for purposes of performance evaluations. Though PDSA was not explicitly tied to Woodson’s common assessments, some departments regarded PDSA as the interim process of formative evaluation used to meet learning objectives addressed by the biannual common assessment. Faculty had relied heavily on their university partner for training in the methods of improvement science, and for one department, this had entailed intensive coaching for a full academic year. At the time this study was conducted, however, departments were expected to implement PDSA cycles independently, and for most this was their first year attempting to do so. By design, the research questions guiding PDSA cycles, the data that were collected, and the study of those data were completely teacher-executed. Woodson had mindfully built some infrastructure to support these efforts, including the designation of weekly staff meeting time to meet and discuss PDSA progress, as well as non-teaching staff to attend and contribute to those meetings. However, the ways in which PDSA cycles were created and implemented in classrooms, and the type of data collected and analyzed, were at the discretion of every teacher.

A short disclaimer is necessary to make clear that the discussion of PDSA within this study is not intended to be an evaluation of the program nor to assess its implementation. The excerpted views expressed by Woodson's faculty with respect to PDSA are also not provided as a representation of their overall estimation of the initiative which continues to evolve. Rather, the perspectives captured here are meant to convey initial teacher reactions and reflections in the first stages of PDSA implementation which inevitably entails processes of adjustment and acclimation to new data use processes.

### **Facilitating Constructive Conversations About Instruction Around Data**

The expectation that teachers use data to inform their instruction suggests, to some degree, that teachers should essentially act as experimental scientists. Although there is certainly a focus on the individual teacher to carry out experimentation within their own classroom, there is also an element of departmental cohesion that can result from constructive conversations around strategically collected data. For the Science Department in particular, the PDSA initiative has been an opportunity for teachers to actively engage in class-based data in ways that are guided by the scientific method and contributive to departmental decision-making. Mr. Macon, a teacher within the Science Department, suggested that his involvement in both the PDSA and common assessment initiatives is valuable because it emphasizes data use as an interactive process of data creation and interpretation:

*You know, I think, personally, as an educator, I think I've... I've learned A LOT. I've progressed... because of looking at data, and writing my own assessments, and... Looking year after year what we need as a department. I didn't have that opportunity before. It was mostly okay, do your thing, go to your classroom, and that's that. At the end of the year, let's look and see how you did with the state assessments, and we'll talk about it. You know?*

*Versus here, we do get... we have a goal in mind. And we get to see what we want to achieve by the end of the year, and the most important piece of course, is the conversation and the data that we're compiling throughout this process.*

When asked which of these elements he thought were essential to his personal practice, Mr. Macon thought for a moment then replied:

*Mmm... hmh. I think for me... Trusting each other. Right? Trusting each other that we have to... We all have a goal. We all have a common goal, and we all need to set up some sort of personal goals to get there. And we need to trust each other to... to achieve what we want to achieve.*

Mr. Macon underscored the importance of dialogue around the data he and his colleagues are collecting. What had both driven and resulted from these dialogues, figured Mr. Macon, was a level of professional trust. Not only had his department been able to establish a common vision, but they also learned to depend upon one another to accomplish curricular goals within their individual classrooms. Departmental discussions around what data were being collected, and consensus as to what those data measure, gave credibility to the PDSA process. In coming together, Mr. Macon and his colleagues are able to rely on the data they have independently collected on their students' progress to monitor, track, and revise instructional strategies as a department.

Mr. Macon explained that his experience using data at Woodson was far removed from the culture of data use at his former school. At his former school (a conventional high school), student achievement data were completely divorced from his instructional practices. He was expected to plan and deliver his classroom content completely independently, and at the end of the year state assessment results were used as a barometer of his effectiveness. In this context, the assessment results held little to no meaning for Mr. Macon, in part because there seemed to be a lack of connection between what the exam measured and his own teaching and learning

strategies. In comparison, one of the aspects Mr. Macon found most valuable in the PDSA process at Woodson was the opportunity to have a discussion with his colleagues about the ways in which class-based data feed into their mutual goals and objectives. The opportunity to establish a cooperative instructional strategy and collectively determine the metrics employed to measure student progress has facilitated constructive conversations about his department's curricular content and instruction.

For Mr. Macon, here the collaborative examination of individual classroom data is the foundation against which departmental goals are monitored and negotiated. In this way, Mr. Macon does not feel isolated in his experimentation, but is reminded that his own efforts contribute to a larger purpose. In the same vein, he is responsible for bearing out the goals and objectives established by his department. Mr. Macon pointed out that implicit within classroom experimentation and department-level reviews of data is a relationship of trust established amongst his colleagues. Mr. Macon's personal use of data is embedded in a process of deliberation that exposes his trials, successes, and challenges in the company of his peers. The Science Department teachers serve not only as an informed audience for one another, but they also share the responsibility of upholding standards of student achievement through their own PDSA work.

It is perhaps no surprise that the Science Department's cultural perspective on experimental science is in step with the tenets of Improvement Science. Mr. Macon talked about the Science Department's involvement in the PDSA process as a "study, if you want to call it that," that his department wants to "last a long time" and "track to see any kind of changes" in student performance resulting from changes in instructional approach. "I mean, that's what research is right?" he asked rhetorically. "Instead of just trying something and changing it next

year, and then in the next year again... it doesn't work.” His emphasis on staying the course and making a concerted effort to actively study the impact of instruction on students’ learning in science is, he believed, “within our nature, you know as science teachers. We make observations, I mean it’s the scientific method. We just follow it. This is what we do every day so it’s... very regimented in terms of that.”

But even with this dedicated, nearly ingrained orientation to methodology, Mr. Macon acknowledged that the PDSA process had not been without its challenges within the Science Department. These were felt most acutely as the Department transitioned out of an intensive coaching program provided by Woodson’s university partner and began implementing PDSA cycles without in-class support. Without this external resource, the data collected by teachers within the Science Department seemed to have taken on a different meaning. Mr. Macon commented:

*There's a lot of things that are going on at school that make it difficult for us to analyze... That was easier when we had somebody to sort of push us. [The graduate students were] always dialoguing with us and saying oh, this would work better, or this is what you could do. So of course, we're going to take that, internalize it, and since we're trying things for the first time, we want to be very good students and start reflecting.*

*But you know, the following year comes along and this is where we are at this point. We're not getting that same support, and we're finding it difficult to reflect and inform ourselves, and to come up with... You know, a PDSA cycle the way we want them to be.*

For Mr. Macon, the process of data collection and documentation seemed more difficult to sustain than was the “study” portion of the PDSA cycle. Finding the time and space to analyze and reflect on data had been difficult in concert with other teaching duties and school responsibilities. As a result, student performance data had become small piles of outstanding

tasks rather than instantaneous instructional insights. This character of the data stands in contrast to Mr. Macon's experience the previous year when the presence of a graduate student "pushed" him into dialogue around his data as a constant momentum in plotting next strategies. This coaching, as well as his personal desire to be a "good student" and learn from the newly-introduced PDSA process, compelled Mr. Macon to conduct PDSA cycles from start to finish. Without this external motivation, however, Mr. Macon was finding trouble "studying" the outcomes of his strategic planning and subsequent data collection. This seems to be true for the Department as a whole and, as a result, Mr. Macon did not believe he and his colleagues were implementing the PDSA cycles with completeness.

In fact, reflecting on one of his instructional goals for the year, Mr. Macon identified in-class support as a primary source of information in improving his students' learning outcomes:

*My goal was definitely, you know, to have them score, in terms of the [common] assessment... at least for 50% of my population to score a three or better. However, you know, I think the fact that we didn't have... much support compared to last year with the graduate students, I think that's what MADE the difference. The more help you can get from an outside [member]... the better, I think.*

When asked what kind of support he felt was really critical, Mr. Macon responded:

*You know, I think it's just feedback. The feedback – communication. Because it does keep you... in a sense, responsible. Oh, I have to reply to this email, I have a conversation with this person, and... you know, the support comes IN that way. And you're able to make changes as you have those conversations. Versus not having any of those conversations at ALL and... You know, having a lot of other things that you have to do.*

For Mr. Macon, the graduate students assisting with the implementation of the PDSA initiative weren't just an additional resource. Rather, they served as an outside eye to his classroom instruction. He found that the data generated by the graduate students – their own observations and feedback – provided value above and beyond his own perceptions and

conclusions. And again, Mr. Macon underscored the advantage of being engaged in conversation about his instruction by someone to whom he is externally accountable. This feeling of responsibility toward “someone else” provides an added layer of accountability not necessarily sustained by self-regulated discipline. Obligated to conversations about his own PDSA process, Mr. Macon not only found that such discussions served as time to reflect on his PDSA work, but that they also flagged concepts and issues to which he would need to return. In this way, his classroom data became a public record of his efforts. With the graduate students now gone, so too was a certain feeling of accountability to a body outside of Woodson. Mr. Macon now found that reflection on his PDSA data was frequently lost amongst other, pressing school duties.

Interviews with Mr. Macon highlighted several different views toward data as it was understood through Woodson’s PDSA initiative. Teacher-facilitated data collection designed around classroom activities has been meaningful as a measure of teaching and learning progress, the basis of departmental dialogue and collaborative planning, and a centerpiece in establishing trust and accountability among colleagues. PDSA data includes not only student achievement results, but also the feedback provided between colleagues and by graduate students serving as coaches. Without the external input of these graduate students, suggested Mr. Macon, not only was an important source of data missing, but also his own data lost an aspect of external accountability. Data that are “unaccounted for” can mean they remain un-studied and, in effect, inoperative.

### **The Utility of PDSA Questioned as an Endless Cycle of Data Collection**

Not all teachers at Woodson viewed the PDSA process as useful as did Mr. Macon, however. This can be partially attributed to the perspective that data collection requirements

were not practically aligned with the classroom-based activities. Ms. Lovell described her early frustration with how classroom-based data collected in the course of the PDSA process can seem overwhelming for a teacher:

*Because here's what you're supposed to do, you're supposed to like... do a strategy right? And then you're supposed to RECORD how you did that strategy. And then... you're supposed to... collect student outcome whatever. And then you're supposed to analyze that. And you're supposed to do it again.*

*And so, I'm collecting two things: I'm collecting that I actually did it, and I'm collecting that.... It's just too much. And then you're supposed to HOLD all of these papers, and then you're supposed to like tally it up so I have a BAR GRAPH.*

*I'm just like, I don't need a bar graph to tell me that I had 80 of my kids... I mean (laughing and raising voice in exasperation), I just don't think that is REAL LIFE....*

*Because how do you record [that students] made an instructional MOVE based on a very intuitive observation that I made? Unless if I keep tally marks, which I guess you can, but then I don't want to collect those tally marks and then put it into like a computer. It makes me mad! (Laughing) Aaahhh (sticks her tongue out and makes "yuck" face). I get really frustrated!...*

*So I guess I see paradoxical things because I want to be intentional in my observations, and I want to keep more, better records... But I don't want to be crazy about it. I don't know, I feel like, we're not scientists, we're teachers. You know?*

In this moment, Ms. Lovell expressed a feeling of exasperation over the burden of data collection required by the PDSA cycles. She was overwhelmed by the need to pedantically document her teaching strategies, as well as devise ways to collect data on how well her students are doing. In addition to the regular demands of teaching, she was finding difficulty in organizing all of these bits of data, entering the data, and then synthesizing them in a (necessarily) quantitative analysis.

Moreover, she questioned whether such efforts actually result in a causal relationship between the data she collects and her teaching strategies. Capturing students' in-class reactions to changes in her instructional approach, she argued, is not easily accomplished in the moment of instruction. This is not just a matter of manual effort. Rather, to do so would require re-orienting her perspective from "teacher," wherein she observes student responses with professional intuition, to "scientist," wherein she must observe her class with a level of objective scrutiny. While not an impossible task – Ms. Lovell suggested she could perhaps use tally marks – there is something strangely artificial about converting her otherwise intuitive observations into a "bar graph" in order to illustrate how her students performed on a singular task. For all of the effort, Ms. Lovell does not entirely see how this type of approach would present results more valuable than what she may have gleaned through her "teacher" lens, however undocumented. Although she personally strives to observe her own instructional practices more closely, and recognizes the benefits of doing so in a methodical, evidenced way, Ms. Lovell weighs this against the practicality of intensively studying her practice through the examination of voluminous data. This, she believes, is the job of the scientist, not the substance of being a teacher.

Because the PDSA initiative requires concerted effort on the part of the teacher to consistently reconfigure his or her approach to instruction, it is reasonable to expect that teachers may encounter some initial difficulty building the PDSA cycles into their already filled days. But this seems to be more than just an issue of finding time or maintaining diligence. As Ms. Lovell explained, constant data identification and collection is an issue of changing culture and personal habit.

*I think to me the whole point about doing the PDSA cycles is for, eventually to be like, almost ingrained in people's rhythms too, so that people are naturally doing it without.... you know that's the whole point. It's like people can internally just be*

*like, how do I know if that's working? Let's collect data on it. So... I think that's going to take time. But I think that's the WHOLE POINT. But I think that we do need more [professional development] around it.*

Ms. Lovell pointed out that effective use of the PDSA cycles require “ingraining” PDSA “habits of mind” into one’s instinctive rhythm. Teachers must be able to fluently associate changes in instruction with the data required to evidence its effects in answer to the question, “What do my students know and how do I know that?” But, as Ms. Lovell pointed out, more professional development will be needed to reach this state of fluidity. She suggested training on how to integrate data identification and collection into teacher practice.

*So I think it would be... having teachers have a better understanding of what is considered. Like what is the data that they could collect that could help us change our instruction—like really make big impact?*

*...What are those data? What does it look like? How can we collect it? And how can we SHARE it? You know, like how can we share it in a way that's not so overbearing, that I have to then type in 20 pages of observational notes. Like, I don't ever want to do that.*

*So I feel like, if there is a really FLUID way where teachers can really do that in a dynamic way, then I feel like that could be really fun.*

Here, Ms. Lovell emphasizes the need to align scientific approaches toward data collection with the needs and capacity of the teacher. She explores the possibility of a closer assimilation between data currently collected by teachers as part of their instruction and robust methods of data collection required for purposes of research. Rather than take on supplementary data collection responsibilities or restrict the type of data teachers collect to align with research-focused rather than instructionally-focused measures, Ms. Lovell wonders if there is a way that teachers could better capitalize on their current data collection efforts. If teachers had a better understanding as to the variety of *feasible* data collection and analysis methods that make sense

within their instruction, it would be a primary impetus for more accurate, relevant data collection in the classroom, and introduce a level of motivational dynamism to the PDSA process. Understanding what varieties of data are considered “credible” in both classroom and research settings would enable faculty to expand their experimental horizons beyond what the template bar graph could represent.

### **When a Focus on Data Use Trumps Good Instruction**

As one Woodson teacher (who wishes remain anonymous) explained, the prioritization of research over practice can have negative effects. She brings cautionary awareness to the potential pitfalls of employing PDSA processes wherein an understanding of credible data is less certain. In this example, she explained how limitations in teachers’ understanding of how to collect meaningful data have inhibited instructional strategies:

*Okay so here's an example. So [our department] messed up and we had this focus where we're going to like focus on student feedback. Which... is easy to measure because... we were going to do like written feedback. That's really easy to measure, you know?*

*But that was TOTALLY not what we needed as a department. Like we needed someone to help us with INSTRUCTIONAL PRACTICE. And implementation of the curriculum.*

When asked why the Department chose to focus on student feedback to begin with, the teacher answered:

*I don't know. I think the PDSA cycle lends people to picking strategies that are very like... tangible. Like, paper (holding up a piece of paper.) And then I can grade it.*

*But... what we really needed is for us to grow in our understanding of the curriculum and implementing strategies to [improve critical thinking that students use during small group discussion].*

*How would you do a PDSA cycle on [group discussion]? OK, like how do I do that? So the instructional move I'm going to make, is I'm gonna... ask more open ended questions. Okay, so that's the change I'm going to make. What kind of data would I gather? As a teacher, what kind of...*

*Okay, so I can maybe keep a tally mark, I think that's the MOST I can do, probably keep a tally mark of EVERY kid who responds to see... whatever. But anything beyond that, I was like I CAN'T do. But that's the instructional strategy that we need to focus on the MOST. But because it's hard to collect data, then we don't PICK that as our group focus. We pick something that's like... so silly.*

Here, this teacher illustrates how her department's PDSA strategy has been dictated by a practical definition of what is "measurable." In many ways, the selection of a measure is based on what is readily "tangible" and fairly easy for a teacher to collect in the course of his/her everyday instruction. And to a certain degree, what is "tangible" is thought of as something easily quantifiable. But she argues that there are more complex instructional strategies employed by her department that should be the primary target. Because she and her colleagues do not readily know how to identify and gather the data that could be used to evidence changes affected by complex instructional strategies, they opted not to select "student discussion" as a group focus despite its importance to the department.

### **Woodson's Identity Crisis**

Woodson's vision of using student and classroom data to inform instructional and strategic planning has come a long way with the introduction of the PDSA and common assessment initiatives. But despite the growing strength of Woodson's organizational data culture, there remains an inevitable degree of variation among its faculty by way of understanding the purpose, process, and benefits of being a data-driven school.

This has been, in some ways, challenging for Dr. Baher, who serves as a primary link between Woodson’s faculty and their university partners guiding the PDSA initiative:

*I know [some of the lower grades] had a HUGE problem and issue with trying to get traction, and they've always [thought this] process is sort of like this mysterious, we're not sure what THEY want, we don't really know what to do, can't someone just come in and help us? Maybe if Dr. Baher could come and run our data it would be okay, right? Like people have that sense of, you need someone else to help you with it. And that's something I struggle with at the school.*

Dr. Baher’s own struggle stems from her reputation at Woodson as its “data guru” and lead researcher. Her comment reflects some frustration with the assumption among some of Woodson’s teachers that PDSA work cannot be done without outside assistance. She conveys that some teachers are thrown by the need to meet “mysterious” external expectations and that data analysis must be so complex as to require research assistance. With a focus on the teacher as the primary agent responsible for conducting PDSA cycles, these expectations run contrary to the tenants of the PDSA initiative. In this sense, Dr. Baher pointed out teachers’ central misunderstanding of the initiative. On the other hand, the needs apparently expressed by these teachers simultaneously reflect the struggle teachers face in transitioning into researchers. What data should be collected, the ways in which data should be collected, and how these data should be analyzed and interpreted call for a greater base of knowledge, skill, and guided practice before teachers will be ready to seamlessly integrate such processes into habits of practice. A culture for data use must necessarily start small and build gradually.

Ms. Lovell, herself an advocate for the use of data to inform school-based decisions, recognized a division among staff in their position toward data. She viewed teacher orientation, in part, as an artifact of personal interest:

*I think my observation in THIS faculty, is that there are... definitely different... [levels of] interest, and engagement around paying attention to data and information to improve my practice.... [There] are teachers that... they just do this, almost naturally. Like, it's really exciting to them, right? And so they're DRAWN to it.... There are other teachers on the other end of the spectrum that SHUT down. They just shut down.*

Importantly, such variation among faculty, while not unnoticed, is not considered a contentious point of rivalry – a dichotomous orientation indicative of whether “you’re in” or “you’re out.” With a deeper sense of nuance, Ms. Lovell paints a picture of faculty who are generally “open-minded” about data use but who also lack a perception of what that exactly encompasses. Data utilization, then, is not only dependent on teachers’ natural inclination toward data, but also their actual understanding and capacity in using data. The latter quality is inextricably tied to the former. As such, Woodson’s “culture of data” is a constant interplay among teachers who differentially locate themselves along spectrums of data use capacity and interest. Ms. Lovell described this culture as something of an “identity crisis.” When asked what she thought the culture of data was at Woodson, Ms. Lovell replied:

*Um, I think we're like, it has an identity crisis.... I think people are confused by it. And I think people don't know... Like, there's all these questions, what do we collect, why do we collect it, and then once we collect it, how do we use it?*

*And then you know, [Dr. Baher] has been really gracious in trying to explain it, you know? But... I don't know, I think it's hard. And I think it's like people don't know, and I don't know, like how hard it is to collect data. I mean, teachers know how hard it is to collect data, because you know how hard it is to grade papers, right? But then they don't know how hard it is to like... Give out a survey, or to...*

*So I think it's, like, it's in an identity crisis and I think what we struggle with is what everyone probably struggles with. I mean, Dr. Baher has repeatedly said to me... You ask questions that you're genuinely curious about. You don't ask a question to like prove a point, right? Because that's not how research is done.*

*I mean, [laughs] but... I think that's how a lot of research is done, you know?... And she's always trying to teach us like, you look at data and we try to first observe, and then we try to analyze. But we're like awful at that, you know? Just like analyze immediately, right? Anyways so things like that, we don't really know how to use it.*

Importantly, Ms. Lovell makes a distinction between teacher “buy-in” wherein teachers are consensually onboard with the data-based activities promoted by the school, and a rooted understanding of what those processes entail and how they are implemented. While the teachers at Woodson generally agree that the use of data in instruction and instructional planning makes sense, they lack a solid understanding of components essential to carrying out these activities in a meaningful manner. The estimated benefits of data-based activities are, therefore, attenuated if teachers go through the motions of data identification, collection, analysis, and interpretation without quite knowing how to navigate those processes independently.

Ms. Lovell also pointed out that developing this kind of teacher capacity is not simply resolved by introductory training and guidance, even if by a researcher. Rather, this seems to be an issue of acquiring a more specific, technical skill set. Teachers are familiar with the difficulties inherent in some types of data collection, such as issuing assignments and grades, but this does not directly translate into the ability to facilitate a strong survey, for example, or to foresee the challenges more commonly encountered in survey administration (such as scale development, survey length, digital vs. paper and pencil formatting, sample selection, etc.). As another example, Ms. Lovell pointed out that specific techniques were also required in the interpretation of data, and she suggested that teachers’ contextual knowledge led more often to assumptions about what the data imply as opposed to observational comments on data trends and patterns.

For Ms. Lovell, an organic cycle of inquiry and follow-up research promoted by Dr. Baher seems far removed from what she guesses is a more common style of investigation in practice – agenda-driven research conducted to prove a point. To satisfy “genuine curiosity” through data and information gathering is perhaps a noble pursuit, but this style of investigation is not inherently woven into the fabric of teacher practice. This is perhaps where the propensity toward data use is seen to play a central role in motivating teachers to pose questions answerable with data. Ms. Lovell suggested that the “identity crisis” encountered by Woodson is probably not unlike the struggle of other schools: data-based decision-making is not necessarily a self-propelling process, and meaningful data use requires as much technical capacity from teachers as it does mindfulness and will.

In his own depiction of Woodson’s data culture, another teacher, Mr. Urbina, echoed Ms. Lovell’s emphasis on the importance of teachers’ direct interaction with data use over and above a general accord to endorse data use practices:

*Because we’re a pilot school, at the beginning we were told we had a lot of autonomy over much of our data. And so, from the beginning you got people who were interested, who know that... that data can be very useful. But that... it’s only useful if teachers and the school community are playing an active role in the tools you’re using to gather the data and then analyzing the data.*

For Mr. Urbina, the utility of data is pegged to the direct involvement of teachers and the school community in its collection and analysis. While he did not go as far as claiming that Woodson had an “identity crisis” in terms of teachers’ capacity to actively engage with data, his perspective paralleled that of Ms. Lovell’s with respect to the need for the intensive resource and capacity investments required to use data meaningfully in their school context. He commented:

*So... are we going to really invest the resources, you know? I feel like the elephant in the room is that this stuff is a lot more complicated than ANYONE*

*thinks. And... figuring out... what it is really what we want to assess with kids, what it is our pedagogy is really addressing, that's going to take a SIGNIFICANTLY greater... contribution of... financial resources to public education.*

*It means smaller class sizes, it means... opportunities for teachers to meet without students and meet together amongst colleagues, and have a coach that is helping them analyze, work with universities... like... graduate students who... have expertise in analyzing data.... Until we get there I think we're going to be spinning our wheels a little bit I think.*

Mr. Urbina provides just a short list of resources he believed were necessary for Woodson teachers to draw on in developing their capacity for data use, none of them simple inputs. Out-of-class time to meet with colleague to sift through and understand how student performance data connect to changes in instructional strategies is imperative but expensive. Teachers' analysis of classroom-collected data would ideally require close and consistent coaching provided by technical experts. The degree to which data can be used to influence individual student progress and performance is naturally limited by the number of students each teacher is meant to monitor.

In commenting on the complicated nature of data use processes, Mr. Urbina – like his colleague, Ms. Lovell - highlighted yet another important facet of the technical complexity in data-based decision-making: aligning data collection activities with instructional strategy. He acknowledged that data identification, collection, and analysis require not only substantial research and evaluation capacity on the part of teachers, but also that being able to integrate data into instruction necessitates careful adjustments on the part of the professional teacher. Curricula tailored to the specific needs of students based on data findings demand concerted teacher reflection on the connections between what and how material is taught to students, and the ways in which student skills and knowledge are subsequently assessed. As further detailed in Part III,

the process of defining, measuring, and recalibrating instruction to address student achievement is both intensive and iterative – a reality Mr. Urbina highlighted as being rarely recognized by proponents of data use in schools.

Woodson College Prep regularly collects, analyzes, and reports student, teacher, and school performance data. These are not just activities conducted at an administrative level or by external researchers; the ongoing institutionalization of data-focused activities, like the PDSA initiative and common assessments, ensure that all of Woodson’s teachers are engaged with data on a personal level. Paradigmatic shifts in the ways Woodson will use data as an influence on instructional strategy is equally reliant on a reorientation toward data use by individual teachers. This is accomplished neither swiftly nor easily. Teachers still struggle in making classroom-based data collection manageable, interpretable, and robust in the eyes of researchers. It is suggested that a better understanding of what feasible data collection could be conducted in classrooms is needed, and that a definition of the criteria of robust research would go a long way toward promoting data use in classroom settings. External support and incentives to analyze classroom data and determine their implications on practice are also viewed to be of great benefit to teachers. Overall, Woodson’s teachers recognize that the integration of data use routines into classroom instruction is not an easy feat, and one that requires a significant investment of resources to conduct in a way that is meaningful for teachers and students.

### **Cross-Case Insights**

Experiences with school-based data and its use in decision-making from both Belleworth and Woodson College Prep highlight an important distinction between organizational and individual orientations toward data use. On both campuses, school leadership, including

principals and teacher leads, has largely endorsed the use of student performance data to guide programmatic and instructional support around specific student needs. The use of student data to inform decisions around curriculum and instruction, however, has also relied on the expansion of this support among all individual faculty members. In Belleworth's case, this meant creating a personal connection between teachers and the meaning of student pass rates in the context of their own grading practices. At Woodson, teachers are expected to regularly engage in data use processes through their participation in the PDSA and common assessment initiatives. These approaches emphasize the necessity of instilling a sense of teacher "ownership" over school data, and the ways in which data are used for purposes of decision-making.

Some teachers at Woodson, however, pointed out important nuances of "ownership" over data and data use routines. Whereas "buy-in" is regarded as the general endorsement and expressed value for data use processes, "ownership" is viewed as the uptake of data use processes in ways that reinforce personal responsibility to the data (and as will later be seen in Part III, an extreme sense of "ownership" can also translate into the exclusion of outside involvement in data use processes in the name of sole proprietorship). While teachers may generally understand the data used in decision-making processes, and though they may broadly endorse the concept of data use to make strategic and instructional planning decisions, there exist substantial influences on data use in the day-to-day context of teaching. The acceptance of data use is not the same as teachers' actual use of data. Rather, teachers' ability to use data in meaningful ways is influenced by wide variations in technical capacity (e.g., the ability to identify appropriate research questions, expertise in measurement, assessment and evaluation, and experience in reading data for emergent patterns and trends), resource availability (e.g., out-

of-class time, coaching, and smaller class sizes), as well as teachers' self-identification as researchers.

In a deeper exploration of what a teacher-led initiative to collect and use student data looks like, Part II details teachers' experiences creating and implementing school-created student assessments at Woodson College Prep. As teachers reflected on their experiences with the common assessment, they provided essential context as to what effective data use in schools could entail.

## **Part II: Data Use in Assessment and Instruction at Woodson College Preparatory School**

Part I discussed several factors influencing data use in processes of decision-making at Belleworth School of Arts and Technology and Woodson College Preparatory School. It was found that one key component of effective data use is the ability of teachers to establish connections between school- and student-level data and their own personal teaching practices. This section shows that, while data use is frequently considered one stage of a cycle – beginning with data identification, continuing with data collection, analysis, and interpretation, and ending in use – data use is best enabled through support at every stage.

The experience of Woodson faculty in developing and implementing school-developed student assessments again highlight the importance of establishing an association between classroom instruction and school data. The strength of this association, however, is very much dependent on teachers' level of involvement in test design, scoring, and analysis. Direct participation in these processes are seen to be essential in reinforcing teachers' understanding of how assessments align with curricular content, and affirm learning objectives as well as how changes in instruction might influence student achievement. Real practice and experience in

developing, administering, and analyzing test results present needed opportunities for teachers to interact with student achievement data and to consider how results can be meaningfully translated into instructional change.

### **The Common Assessments**

In an exercise of Woodson's autonomy over assessment, the teachers at Woodson College Prep have, for the past several years, been focused on developing and implementing intensive subject-based student assessments as a way of measuring student performance at the beginning and end of each academic year. Termed "common assessments," these school-owned exams are used in place of the standardized "periodic assessments" facilitated by the District. Woodson's approach to the common assessments stem not just from a desire to depart from the District's assessment of student performance (tests at one point boycotted by the teachers' union), but rather the determination to "ground people's sense of ownership over the measures that would be used to gauge their progress."

Each department within Woodson's upper school oversees the management of its own common assessment including its content, facilitation, scoring, and results analysis. Recognizing the need for technical support and capacity building in these skill areas, teachers have capitalized on Woodson's relationship with its university partner in developing and adapting test items, identifying and implementing scoring criteria, and conducting analyses of results for the purpose of informing instruction. In addition to finding funding (e.g., grants), to engage its university partner in test development, Woodson has also allocated a substantial portion of its budget to release days for teachers, affording faculty the opportunity to work collaboratively on the development of the common assessments and the review of biannual data.

While Woodson has made an institutional commitment to the common assessments, each department's experience designing and implementing their assessment is unique. This chapter details the various ways in which Woodson's departments have each considered the content of the exams, the ways in which they are evaluated, and how student performance data have been used to make instructional decisions. Collectively, these accounts suggest that the more regularly teachers engage in the many different stages of test construction, facilitation, and review, the more likely they are to make meaningful use of the resulting data.

### **The English Department: Assessments and “The Hidden Curriculum”**

Through his own work with Woodson's university partner, Mr. Urbina was introduced to a writing assessment designed to gauge the academic writing proficiency of incoming university students. Mr. Urbina found himself drawn to the assessment for its use of a comprehensive “continuum” measuring students' writing skills. Rather than focusing on student deficiencies (i.e., “What students *aren't* doing”), Mr. Urbina recognized the orientation of the continuum on student ability (i.e., “What are students doing *right?*”), as well as the test's explicit inclusion of student voice (i.e., “Here is what the expert thinks of this passage, what do YOU think?”). These key features, from Mr. Urbina's perspective, are also what piqued his colleagues' interest in using the university writing exam as the basis for the English Department's common assessment. The Department subsequently worked with Woodson's university partner to adapt the university-based assessment for use at a high school level.

Mr. Urbina highlighted how the adaptation of the assessment and a collaborative commitment to the criteria introduced by its continuum influenced the English Department's instructional strategy:

*It was several pieces. One was getting teachers to kind of wrap their heads around the assessment itself and like figure out... I mean, essentially, with every assessment... what... the hidden curriculum is.... What exactly is this assessment asking my students to do? And what do I know what my students can or can't do, and what do I currently teach? What does my instruction design do or not do?*

*So that in itself can be sort of an orientation moment... for a lot of teachers, I think. And... luckily it wasn't a BIG jump from what teachers were already doing. But it was definitely... a moment where teachers were like... "Oh ok, so this is what we're saying we want to be able to do."*

*And our sixth grade teacher... So HE is in a place where he's seeing... what various preparations [students] are getting in elementary school, and then what they're headed towards in high school. So he was seeing a lot of preparation around fictional writing and a lot of personal narrative [in elementary school]. And then, he saw this, our assessment, and he's like, this is really non-fiction based and it's analytical. Uh... so he had to do a little bit... of orientating.*

Mr. Urbina explains that, while this external assessment served as an important starting point and his colleagues found its scoring continuum particularly meaningful, assessment adaptation required some significant re-orientation of the Department's approach to writing. Indeed, "orientation" seems almost an insufficient description of teachers' process of unpacking the "hidden curriculum" embedded within the assessment. If the assessment was to be used as measurement of student writing capacity, teachers needed to ensure that their own instructional content aligned against those measures. Through his words, Mr. Urbina walks through some of the lynchpin questions associated with understanding how the assessment caused a reevaluation of current instruction: What exactly is this assessment asking my students to do? What do I know what my students can or can't do, and what do I currently teach? What does my current instructional design do or not do?

For most of his departmental colleagues, Mr. Urbina suggested that the divide between classroom instruction and those aspects measured by the assessment was not too wide, and that

subsequent adjustments were not terribly laborious. He did point out, however, that the Department's focus on analytical writing presented some challenges in bridging curriculum between the upper and the lower school, the latter of which focused more on narrative and fiction-based writing. Transitioning students from one form of writing to the other thus became an important focal point for the middle school grades in preparation for the conversion in assessment content.

As Mr. Urbina detailed the rollout of the exam over subsequent years, however, it is clear that the instructional shifts triggered by results of the common assessment had been no small undertaking. To start, the English Department invested a year or two working with Woodson's partner university to level the college-based assessment for use with high school students. Following this accomplishment, a first round of the common assessment's implementation revealed that the Department's teaching on writing was "pretty strong, and kids understood... how to structure an essay, and transitions, and how to insert evidence, but their problem was that... they were hitting, essentially, a glass ceiling with their reading, and understanding the arguments, and being able to... pull out argument from the reading." The Department next decided to "attack low hanging fruit" by focusing on student annotation and "chunking the text" rather than allowing students to skim through the reading passages. This required a revised assessment format, where physical space – in the form of wide margins alongside reading passages – was created to encourage students to make annotations and take notes as they read.

Following another round of common assessments, the Department was encouraged to see small "bumps" in test performance wherein some student writing was definitely noted to improve. Despite this, teachers continued to observe a "glass ceiling" in score attainment. An analysis of writing samples suggested that students were still misrepresenting text and

experiencing difficulty with reading. Mr. Urbina described how the Department's next area of focus would be to invest significant time and energy into identifying students' individual reading levels, as well as providing in-class libraries organized by reading "lexile scores." The Department is currently working to identify an assessment that can both accurately and efficiently determine students' reading levels, as well as procuring books for leveled reading libraries.

In summary, the efforts undertaken by Woodson's English Department show fairly dramatic instructional shifts in preparation for, and response to, the common assessment. Feedback from the assessments has consistently informed department-wide strategies to reading and writing. It has also been heavily reliant on teacher expertise to translate into instructional moves. On an individual level, for example, Mr. Urbina's reflection on the high-scoring student essays have prompted him to seriously consider dedicating more class time to vocabulary, supporting his students in using the "Charty Graphy" strategy in outlining their arguments before writing, and to more generally strengthen students' identities as readers, writers, and thinkers by creating structured activities which focus on personal voice and narrative.

Given his experience with the common assessment, Mr. Urbina found it impossible to separate meaningful data use in Woodson from direct teacher engagement in data collection and subsequent data analysis. In adapting and developing his department's common assessment, the teacher as expert practitioner has been an essential translational link between skillfully-designed assessments, curricular alignment, and instruction. This is not only important to ensure that class content addresses the student performance standards endorsed by the assessment. But, as curricular approaches form and flex around assessment findings, teacher feedback has been imperative in revising test content and format. For the English Department, implementation of

the common assessments has demanded a substantial amount of time and attention in reviewing test content, collaboratively defining departmental learning objectives, goals, and standards, reading and scoring hundreds of student essays, and converting interpretations of test results into personal changes in instruction. Without this degree of teacher involvement, however, the common assessment would lack relevance to both teachers and students, either failing to measure prioritized constructs and skills, or producing results that teachers would not readily know how to contextualize. The English Department's use of the common assessments findings therefore relies on teachers' explicit interaction with test content, format, and scoring processes.

Although the English Department has gone through several iterations of its common assessment, it appears that faculty are generally pleased with its form and substance. However, for other departments, such as science, progressively unpacking the common assessment's "hidden curriculum" has led to the realization that a more significant investment on the part of teachers will be need to configure an assessment well-fitted to their expectations of student performance.

### **The Science Department: Aligning Standards, Measures, and Instruction**

The Science Department has worked steadily with Woodson's university partner to develop its common assessment, initially piloting the University's ready-made science examinations whole parcel. Over three years, this partnership worked cooperatively to modify and tailor the assessments to the Science Department's subject content and to create a forum for students to exhibit their knowledge through expository writing. Mr. Macon found these recent versions of the common assessment particularly useful in informing his department's approach to

instruction. He explained how his colleagues began to fold their response to assessment results into classroom activities:

*Now that we have our assessment... and what we notice in the assessment – the kids needed help in writing. So what do we do as a department? Oh, we need to elaborate a little bit more on how well students do the laboratory reports. THAT’S our only USEFUL tool that we can sort of help students in writing. And so now our laboratory reports are completely elaborative. They’re Common Core aligned, they were co-developed with [the University], and so that definitely took it to the next level.*

Mr. Macon discusses how student performance on the science common assessments highlighted the need for improvements in writing about science content. Thinking about how they might reinforce writing skills in their own lessons, Mr. Macon and his colleagues focused on the kind of work incorporated into in-class laboratory reports. The past year’s departmental efforts focused on redefining lab report requirements so that students could exercise writing techniques later needed on the common assessment. In turn, teachers found that these assignments were thus aligned not only with the common assessment, but also with Common Core standards addressed by the common assessment. But the Science Department also found that student writing needed to be more closely tracked as a way of promoting progress throughout the year. Mr. Macon continued:

*Late in this year we decided to implement sort of like a... mini assessment..... It’s an assessment basically that we’re trying to create four times a year. Making a laboratory report in order to determine, you know... how much [students] are improving.*

*So... we are sort of now focusing not only on... developing those specific assessments, but we’re also focusing on developing... you know, strategy. Teaching strategy based around those assessments. So for example, we have like a double entry journal that helps out their writing skills, which can LATER be used in the introduction of their laboratory report. We have graphs and charts that we USE periodically that will, again, help out in their laboratory report. We*

*have an ANALYSIS tool, a teaching tool, like [explaining] matter, in order to help students... understand graphs and charts a little bit better—in order to analyze data a little bit better. Those are a few examples.*

Here Mr. Macon describes how the decision to engage students in more consistent assessment activities supports a finer-grained view of student ability, as well as the configuration of classroom instruction, in order to support student success. The “mini-assessments” help the Department shift its focus from the laboratory reports as a final product to working with students progressively through its separate components. Complementary skills-based activities, such as journal writing, the use of charts and graphs, and practice conducting data analysis, are each addressed in modular form and are eventually fed into the final laboratory report. This scaffolded approach to composing laboratory reports, combined with more frequent assessment, allows the Science Department teachers to identify and address student need areas before the year-end administration of the common assessment.

The past year marked still another change to the Science Department’s common assessment process. Although teachers within the Science Department had been trained to score the test, the University conducted this work on behalf of the Department up until this past year. Teachers’ ownership over the scoring process has proven to be an important exercise in data use, albeit in unanticipated ways.

As with the English Department, participant observations of the Science Department’s scoring sessions this year revealed that teachers’ direct involvement in reading and evaluating student responses raised important questions as to whether and how their instruction was reflected in the performance expectations promoted by the exam. Mr. Macon reflected on the group’s takeaways after scoring the assessments with his colleagues this first time:

*So we figured that, you know, it's a lot of the conversation we had during the scoring which was... we need to change the rubric. A lot of the rubric items in there were a little vague. And so, we thought well... what is considered "sometimes," or "always?" So these were like key words there that we are wondering a lot about....*

Mr. Macon suggests here that the process of scoring individual student essays forced each teacher to actively consider the parameters of the assessment's scoring rubric. Where a teacher was unclear as to whether a student's essay earned, for example, a "sometimes" or "always" on a five-degree scale for any given criteria, they would confer with a partner, each taking turns reading the essay. While in many cases, a consensus was reached within pairs, there were certainly gray areas detected wherein a clear answer was not obvious, even when presented to the entire group.

As another example, some discussion arose around what constituted adequate evidence of an "argument" in a student essay, one of the main criteria detailed by the scoring rubric. One of the teachers from the upper school suggested that an argument would involve evidence introduced by a student leading to the presentation of additional (not reiterative) evidence. A counterargument should also be present. However, a middle school teacher suggested that this definition might be grade specific. For Grades 6-8, she suggested, the claim is laid out for the student in the assessment prompt, and students were asked to support that claim with textual evidence. The younger students were not necessarily expected to evaluate the evidence presented by the assessment and select pieces to support an argument of their choice. Hearing all of this, another teacher – visiting from another department – questioned, "Is this still an argument then, or an explanation?" She went on to explain how her own department is using an "explanation rubric" rather than an "argument rubric" for just this reason. While the language differences were

not substantial (often the word “explanation” was simply substituted for “argument”), this would affect the ways in which student essays were considered and scored.

There was also some difficulty in interpreting the “referencing” domain on the scoring rubric. Was it enough, some teachers asked, if a student “alludes” to the reading passages, or did he/she need to be able to provide a specific citation, such as a sentence beginning with, “In the reading...” One faculty member, a former lead teacher currently working outside of the classroom, put forth that Grade 6-8 students should be able to at least flag from where they are drawing ideas that are not their own. Other teachers seemed to take a more general approach, expressing their opinion that paraphrasing or clearly drawing from the ideas presented in the passage would be sufficient. The group agreed that the test prompt was not explicit about the use of citations or references, cuing some deliberation over revising the prompt. The out-of-classroom faculty member suggested that whatever the Science Department decided for its rubric, should be incorporated into teachers’ instruction as well. As she held up the rubric, she explained, “This part of the rubric is the rubric you’d be using all year. You should create a teaching rubric that goes along with this. Have the kids grade their own papers using the rubric, and you all know what you need to do.” There were some audible, “Oohhs,” heard in response from some of the teachers to whom this was a new and intriguing idea.

In both of these examples, the teachers within the Science Department found themselves weighing elements of the rubric not just for the sake of scoring, but to better understand how scoring aligned with their expectations of student performance. Was the test prompt clear about those expectations? How might the same rubric criteria apply across grade levels? What implications did the rubric have on the way students should be prepared for the exam?

In the first example, even slight changes to the rubric's language, such as in converting the word "argument" to "explanation," would substantially impact how a teacher might rate the quality of an essay. Choice of terminology would also be a proclamation of how teachers expect their students to use evidence in their written responses. Careful thought by the entire Department would be required in considering a revision to this single component on the scoring rubric. In the second example, the suggestion to create a "teaching rubric" from the scoring rubric again highlighted acknowledgment of curriculum embedded within the common assessment. The faculty member in this example suggested that whatever the Science Department ultimately decided by way of scoring criteria related to "referencing," these criteria should be incorporated into everyday instruction. Not only should teachers be clear about what constitutes "referencing" in their lessons, but students should also be able to review peer essays and identify whether adequate referencing is present. Her suggestion implies that the scoring rubric is far from a passive element of an external assessment meant simply to observe and detect student skill. Rather, the rubric serves as an open declaration of learning outcomes that are, in turn, practiced and understood by students and teachers throughout the year.

Through this experience of scoring, and deliberately walking through what those scores imply in terms of student performance and teacher instruction, the Science Department determined that it would need to revise its rubric and some of its test prompts. But what this process will look like – when will teachers convene to revise, how the new rubric content will be selected, and what new content should be considered – is still to be determined.

With this new introduction to scoring, the science teachers' use of common assessment data had become multi-faceted. No longer were assessment scores the sole area of focus. Rather, in reviewing the full content of student responses and translating these into rubric-based scores,

teachers began to see the "hidden curriculum" inherent within the assessment. In unpacking these implied standards of performance, faculty were compelled to reflect on how their own classroom instruction adequately prepared students to do well on the assessments and whether the assessments were designed to accurately reflect skills and knowledge focused on in the classroom. While teachers within the Science Department may have previously recognized this connection, it was not until they began to interactively engage in the practice of scoring student essays that they developed a deeper understanding of how their rubric “fit” to classroom content, curriculum, and student work. In some regard, because of some mismatches identified between the test’s instructions, the scoring rubric, and teacher expectations of performance, the Science Department considered the common assessment results somewhat flawed. Nevertheless, the thorough, structured, collaborative review of student responses became an essential source of data in informing needed test revisions and potential changes in instructional approaches.

The last case within this section highlights experiences within the Social Studies Department and details a similar process of assessment administration and scoring for one more teacher. However, unlike the constructive conversations surrounding the common assessment observed in the Science Department, the experience of social studies teacher Ms. Gilman paints a picture of how confusion, frustration, and aggravation can also characterize data use associated with test development and facilitation.

### **The Social Studies Department: Misalignment and Disenchantment**

Developing assessments that feed meaningful data back into instruction is no easy task. But for Woodson’s English Department, as well as its Science Department, this undertaking had been strongly supported by inputs from Woodson’s university partner, as well as an enduring

commitment from departmental faculty to the assessment itself. A case from the Social Studies Department provides a final example of common assessment implementation that is perhaps less directed. It presents a more extreme example of the difficulties inherent in wrestling through the process of data creation and use. It also taps into Ms. Lovell's depiction of Woodson's "identity crisis" wherein some teachers feel that their capacity to identify, collect, analyze, and interpret data falls short of their desire to do so.

Ms. Gilman described the process by which she and her colleagues opted to develop their own common assessment:

*There was a bunch of pressure from the District to do their assessments. And we were like, WE'RE not doing them! Look at these assessments! This is ridiculous! We're not teaching to a test! Duh, duh, duh, duh! (dramatically her shaking head from left to right with each exclamation).*

*And, so then... You know, I think there was... some pressure of like, okay, don't do them, but you have to do something. So... then it was like, okay, we'll create this assessment that does sort of meet our vision.... So... then we tried to create these assessments, which it turns out, they're super hard to do! (Laughing) And yeah, it's been a little bit of a struggle of like... How do we basically create our own data? How do we show that yes, our students are improving and... improving based on what we do? Like if we use a strategy, they actually improve.... So that's what we're deeply entrenched in, and it's really hard. (Laughs)*

As Ms. Gilman detailed, the driver to create common assessments within the Social Studies Department was to develop a measure of student knowledge and skill that more closely reflected the vision of the Department than externally developed District criteria. This would avoid the problem of "teaching to the test," or the need to focus on the irrelevant standards introduced by the District's assessment, which the Social Studies Department felt was lacking in both its exercise of critical thinking and progressive social justice content. With the freedom to develop a teacher-driven assessment, however, Ms. Gilman was expressive about the challenge

this presents to the Department, particularly in evidencing a causal link between instructional strategy and student academic achievement through a well-designed test.

She went on to explain that the Social Studies Department also tapped Woodson's partner university and was able to not select an assessment (ready-made assessments were not available for some of the subjects covered by the Social Studies Department), but a scoring rubric to adapt for the Department's use. In implementation, however, Ms. Gilman was disappointed with these adopted criteria:

*But it turns out their rubric really sucks. At least for what we were trying to do, or in my opinion. But because we used that the first year, there's some sort of pressure to keep using it. Because then, otherwise, HOW do we collect data, or HOW do we show that we're improving?*

*And so... as a teacher that's been doing it now for a couple of years, to tell you the truth, my interest in it has kinda' fizzled. I feel like, I KNOW that certain things we're doing are really helping, but that it's not showing on the stupid rubric. The rubric is dumb. It's not even really like reflecting our goal.*

An air of disenchantment wafts through Ms. Gilman's depiction of the common assessment as she explains the Department's commitment to a scoring rubric she feels does not adequately capture her students' progress. Her own teaching, she noted, does not seem to align well with these criteria. The beneficial instructional moves she believes she is making are not detected by the scoring system. As an example, she explained that the rubric equally weighs the correct use of grammar against all other aspects of a student's essay. In her own professional opinion, however, correct grammar is secondary to whether a student understands and aptly expresses a historical concept. As a Department, she felt that the teachers' goal of enabling students to analyze primary source documents as evidence for their own arguments is overshadowed by a rubric that prioritizes different criteria.

Still, Ms. Gilman feels “pressured” by her department to use the same rubric each year. Dr. Baher, who has been instrumental in guiding the Social Studies Department through the analyses of their annual common assessment data, detailed in a later interview that the lead teacher for the Department opted to continue using the rubric as a matter of “staying the course” and “proving what we have.” The Lead Teacher for the Department decided to “fix” current assessments rather than engage in a search for new test content all together. Dr. Baher emphasized that this was an important leadership decision, particularly given that the Social Studies Department’s assessment strategy was “violating so many assessment best practices” so that the validity of their results was questionable.

The sustained use of the “ill-fitting” rubric, however, seemed only to culminate in Ms. Gilman’s disregard for the common assessment exercise as a whole:

*So kind of what we decided to do this year was... we have to use this whole, sort of same rubric because... that's how we collect data, and isn't the goal to show that our students are improving? So we keep using it, but we're sort of only going to focus on these [particular criteria bands]. And so like whatevs if our scores go down on everything else, we're like, JUST going to pay attention to these ones. (Laughs) But even THEN, I don't know... how do you sort of be... teacher-driven, and create your own [assessments], and make sure they really work AND do the scientific-y showing data?*

*Because it's a big mess. Like the first year... We, like, paid money for these university people to grade the essays. So we did all the work, do the essays, and they graded them. And we did really bad. (Laugh) And so part of us were like WHO ARE THESE PEOPLE? This is a great essay!*

*But I mean, like well, you don't really know who that student is, and where they started. You know, it's so complicated. And NOW, we don't have money to pay those university people anymore. So now we are grading them. But we're using the same rubric, and I'll tell ya' right now... Me using that rubric, I'm all like, “Oh! Five points! They totally nailed it!” You know?! (Laughs while making large checkmarks in the air)*

*It's so far from scientific... this whole thing. Part of me is like... God, we don't even need to use that rubric because it's like MEANINGLESS how we're using it! ... Here we're trying to, sort of create our own data, but... to create the sort of formal data that is supposedly accepted is so hard and it's definitely... using a lot of our time and energy doing this. I mean because you're doing this "trying to prove" thing.*

Here again Ms. Gilman reiterates the technical challenge of producing reliable, “scientific-y” data from a self-created assessment. As her department decided to narrow down its focus to student progress along select rubric criteria, she questions whether it is appropriate to disregard the remainder of the rubric: Is this scientific? Or is this selective view employed for the sake of producing scientific data? She continued, noting that even the professionally-graded common assessments conducted by Woodson’s university partner did not seem to produce results in step with the Social Studies Department’s view of student achievement. Then again, she considered how her own liberal use of the same rubric likely lends itself to score inflation. In all, Ms. Gilman cannot see how the common assessment, meant to “prove” progress in students’ skills and knowledge, can be thought of as “scientific” if the use of its scoring rubric has been used so inconsistently.

Importantly, Ms. Gilman alludes in this passage to an underlying sentiment that the very system of assessment and scoring, especially if conducted outside of the Department, could not possibly capture the complexity of student progress. For Ms. Gilman, “knowing who a student is and where they started,” are fundamental components to understanding student performance. The need to have this close understanding of a class, its individual students, and what substantiates their “progress” feeds into why Ms. Gilman holds teacher-developed assessments – as opposed to externally-created tests – in such high regard. Although Ms. Gilman believes that a teacher’s input is critical to creating a reliable assessment, she found that developing her own test for her

own subject area – one that needs to somehow adhere to the department’s rubric – is much more difficult than she anticipated. In an interview at the end of the year, she commented:

*I think it's different for each department, and each PERSON, and... I don't know why it was so HARD for me. I'm just like, oh my God, I am so frustrated! DEFINITELY a piece what's hard is like...*

*I politically really believe in making our own assessments, and duh, duh, duh, but... You know, truth be told, I spent TONNNNS of time, like HOURS, putting together this assessment that... I really thought sort of met all of the goals. We had all of these meetings talking about it. How should it be? And like... the two new teachers in our department, I KNOW they spent lot of time, even more than ME. The one told me, she was like, I spent at least 10 hours putting together just this one assessment that the students, you know, take an hour and half to take. That's a lot of time.*

*And then... as we're all grading them... Really it turns out that the students did not do well.... And maybe you're like oh, well the test isn't fair, maybe that's why they didn't do well, or didn't adequately measure, or... Really they didn't do well because I DIDN'T PUT IT TOGETHER WELL. And I didn't teach it well, because I didn't understand... what we were doing!*

*So NOW it's like... Here we are all grading my assessment, that I MADE, and basically criticizing that I didn't make it well! And that my students didn't do it well! Yet, I spent all this [explicative] time doing it! And wasn't given like... ahhh!!! So I'm all frustrated.*

By the end of the year, Ms. Gilman is remarkably honest about her stance toward the common assessment. Her experience this year had been nothing short of a struggle, both in developing her own assessment and in wrestling with the outcomes of a flawed design. Ms. Gilman admits having spent a great deal of time talking about the tests in meetings with her colleagues and investing personal resources into creating what she believed to be a test capturing the Department’s goals. However, at the end of the day, she felt that evidence of her students’ capabilities had been undermined by her own deficiencies in test development. Ultimately, Ms.

Gilman attributes her missteps in underpreparing her students for the assessment to a lack of understanding about what the Department is doing all together:

*Because this rubric is what we're going to be judged on. I don't know how to make an assessment with that rubric, and if all this research has been done how it should be... and you can only have four [primary source] documents, not six and... Then FINE! Just gimme' that one! I don't know!*

*I spent all this thought and time and then... they didn't even do well! And it's really because of what I did. And I was even telling them, like, you know, I like gave them LOTS of structure. Like, in your first paragraph you should have this topic sentence, and... maybe too much, but I was just trying to like... I don't know, this is what you have to do for this.*

*And then... I was sort of criticized because... they didn't make an argument, you know? I was just like oh, well, I didn't TELL them to do that. So... the whole thing was very frustrating. And it ended up making me feel very sort of like isolated and frustrated (tearing up). Not bringing our department together. So I don't know. Like uggghhh... glad it's over!*

Ms. Gilman is the first to take responsibility for her students' apparent underperformance on the assessment. But thinking about the shortcoming of her test brings her to tears as she is overwhelmed by feelings of "isolation" and "frustration." She recognizes that the rubric chosen by the Department is an outward statement of performance that she and her students will be "judged on." However, she feels ill-equipped to develop an assessment that aligns with the rubric. For example, despite her emphasis on clearly outlining the requirements of the essay, she seems to have left out the overarching directive of "making an argument." She briefly mentions her surprise at how much the exact number of primary source documents students were expected to evaluate mattered.

Ms. Gilman's side comment about "all this research" linking assessments with rubrics divulges a perspective completely outside of this technical realm. Increasingly aware of the

centrality of such research, Ms. Gilman concedes that it should take a more predominant role in guiding her own assessment development. In fact, seemingly defeated by her own weaknesses, Ms. Gilman was ready to forfeit her philosophical stance on teacher-developed assessments for one that already meets all necessary research design requirements:

*I really do hold this political belief that no, teachers should make their own assessments, because we know our students, and... we know the CONTENT, and we are the ones setting the goals, and... but yeah... if it's going to be like THIS, I feel like... [expletive]. Make it for me, I'LL look at it, I'll make sense of it, show me what they're going to be graded on, and I'll figure out a way... to... teach valuable skills, and... have them do well (laughs).*

The feeling of being lost in the common assessment process is, at this moment, completely demotivating for Ms. Gilman. She would rather have Woodson's university partner develop a test for her than fend for herself at the drawing board again. She later went on to remark on how impressed she was with the English Department's common assessment, even though it hadn't developed the test completely independently. She expresses interest in following a similar process where she could adapt material from an existing exam. But for the time being, Ms. Gilman feels as if her own common assessment is completely "not useful," and is aggravated by how poorly it exhibits her students' aptitude. She takes this setback very personally, and is deflated by the notion that she is potentially the only teacher within her department that feels quite this way.

Although Ms. Gilman may have, at the time, felt alone in both her frustration and feeling of ineptitude in measuring student performance, her perspective is undoubtedly shared by many teachers who grapple with identifying, collecting, analyzing, and interpreting their own student achievement data. As it turns out, the experience of stepping through these processes independently was not enough to produce data considered useful by the Social Studies

Department. On a conceptual level, Ms. Gilman highlighted her own need to find a better understanding of how her department's learning objectives directly map to its assessment activities and its externally-derived scoring rubric. On a technical level, she expresses the need to know how to connect the design and form of her self-developed assessment with the rubric's standards of performance. On a philosophical level, she would like to have a better idea of what kinds of measures could support a constructive view of student capacity rather than simply identifying student "underperformance." Although Ms. Gilman had the opportunity to regularly discuss these issues with her departmental colleagues, the development and trial of her assessment was, ultimately, under her sole purview. Departmental discussions seemed somewhat removed from the actual process of assessment design, and without the chance to pilot items and see how students might respond to her test in advance of full-scale distribution, it was determined only after its year-end administration that her assessment contained substantial flaws. Despite the effort she had invested in creating her assessment, her data were considered useless in their reflection of student ability.

### **Cross-Participant Insights**

The utility of Woodson's common assessments has been discussed in several ways throughout this section. Student achievement data derived from the assessments have certainly fed back into instructional changes as discussed in the examples of Mr. Urbina and Mr. Macon. Their review of student test scores, reflection on student performance through the lens of rubric standards, and their in-depth study of student responses all contributed to identifying areas requiring further pedagogical focus. Classroom activities were either developed or modified to bolster student performance in targeted skill areas, and specific learning strategies were

reinforced as a way of preparing students for testing. This could be viewed as “teaching to the test” but, because the common assessments for the English and Science Departments were a manifestation of prioritized student learning outcomes, and because these learning outcomes were well aligned with teacher expectations, the tests served as both a useful assessment of student skill and a validated benchmark of performance. Work towards improvement in test performance has become equivalent with student improvement in competencies central to classroom content.

As such, meaningful use of the common assessment data not only rely on teachers’ ability to analyze and interpret student results, but is also dependent upon teachers’ active involvement in the process of assessment development, administration, and scoring. This level of engagement is critical in cultivating teachers’ working knowledge of how the assessments connect with, react with, and respond to changes in student performance. Fully understanding these relationships ensures that the results produced by the assessments bear interpretable, actionable meaning in classroom contexts. For example, the Science Department’s participation in reading and scoring student essays this year shows how essential this process has been to fully understanding the implications of rubric criteria, not only on the evaluation of student performance, but about the ways in which students are prepared throughout the year to meet those standards.

The utility of Woodson’s common assessments cannot be fully understood as the result of a unidirectional process of data collection, data analysis, and, then, the translation of results into instructional change. It is not simply the output of the assessments that are important to use. Rather, just as the common assessments have influenced instructional strategy in the English and Science Departments, teachers have also worked to revise their assessments in accordance with

their own needs. Assessment development in these departments has shown that test design must be responsive to format fine-tuning and content adjustment as teachers iteratively improve the ways in which tests elicit student knowledge. As such, trialing and revising items and scoring criteria are essential to teachers' understanding, value, and subsequent use of assessment results.

When learning outcomes are not clearly connected with assessment design, content, and scoring criteria, however, data use is compromised. For example, although Ms. Gilman was an active participant in the design, development, and review of her department's common assessments, she found that involvement in these processes was not sufficient in bolstering her understanding of what makes a "good" assessment or why making a "good" assessment matters. Rather, her experience of the common assessment process revealed that curricular content knowledge and teacher-identified standards of student performance need to be paired with some technical expertise in order to develop a test yielding usable results. Ms. Gilman also identified gaps between the learning expectations she maintained for her own students and those upheld by her department and its assessment scoring rubric. Her philosophical stance toward testing suggests that she has serious doubts about the ability of exams to accurately capture student aptitude. While she questions the appropriateness of the rubric guiding her department's common assessment she feels, all the while, at a loss to produce an alternative instrument that equally upholds the tenets of measurement validity and fair measures of student progress. Ms. Gilman's inability to connect the common assessment with her own perceptions of achievement and credible evidence of student capacity has been debilitating in her design and implementation of a test this year. Despite the great amount of time and effort invested in the common assessment exercise, Ms. Gilman is devastated at signals of her students' underperformance resulting from, what she considers to be her faulty test. Not only does she find the results useless,

but she also feels isolated and frustrated by her shortcomings as a test developer and a lack of collegial camaraderie within her department.

At Woodson, teachers' detachment from processes of assessment development and analysis is associated with their lack of value for, and use of assessment results. Several participants have made a sharp distinction between their value of data resulting from the common assessments and standardized assessments administered by the District, the latter of which they find irrelevant in its measure of student ability. In addition to doubting the credibility of standardized, or "one size fits all" measures, they are some feelings that the ways in which these assessment data are used tend to be more punitive than constructive. As Mr. Macon reflected on his former school's review of state test scores, he explained:

*You know the kind of analysis we got was whatever we received back from... the state. "This is what you received on the [state] scores," and you know, you were sitting in front of the whole school and the Chemistry Department did so poorly. And of course I was the only person in the Chemistry Department.*

*So it wasn't used effectively, and it was more like [being] reprimanded rather than, you know, this is how you can grow, this is what you can do with your data, blah, blah, blah, blah. But, it wasn't handled, I don't think, in a very professional way. It could've been done differently I think.*

Mr. Macon's account is similar to that of other interviewees who explained their experience with assessment data use as a public review of results emphasizing areas of underachievement. Analyses of the data tended to be aggregated at the department level and were not necessarily presented in ways that supported further investigation or actionable next steps. In the case of Mr. Macon's prior experience, not only did he feel that his department was identified as falling behind, but because he was the only teacher in the Chemistry Department, he felt personally criticized. Without colleagues with whom to discuss the results, Mr. Macon was

left on his own to determine how to respond in ways that could improve student achievement – a daunting task at minimum. Regarding teachers as consumers of data rather than agents of data production ignored the role teachers need to play in order to effectively understand and interpret student achievement data as well as to translate findings into instructional improvements.

As a final point, another use of student achievement data still to be approached by Woodson's teachers is the communication of results to students. This is because the ability of students to make sense of their scores in a way that positively supports their academic improvement is believed to require careful framing. Mr. Urbina described his own considerations in delivering common assessment results to his students:

*If a kid asked, I would show them. But I felt like we weren't ready to... we hadn't crafted... talking points, as a department, to... present the data in a way that would be meaningful to kids – that would not be dehumanizing in any way... That would... honor... what they brought to the table as opposed to making them feel inadequate because they hadn't scored a perfect score.*

*I mean, the kids are constantly living in a culture of... data being given to them and REALLY not understanding that context of that data.... And I feel like that happens ALL the time. And... if you're not, as a kid, if you're not getting the top score, the perfect score, then you're a failure. OR, it's like... they're like, "See, nothing's going to change, like, nothing's changed."*

*I wasn't comfortable... asking the Department to kind of give back data to kids unless we... en masse... until we figured out how to make it meaningful for kids and not to make them feel... to honor what they ARE doing and not what they're NOT doing.*

Mr. Urbina's perspective is that students are entitled to a presentation of the common assessment data that "honors" their current capabilities and presents concrete ways for them to grow. Too often, he claimed, students are given data out of context and without much explanation. Left to draw their own conclusions, students tend to interpret scores less than

perfect as a form of “failure.” In Mr. Urbina’s opinion, the English Department needs to first develop a thoughtful approach to data dissemination before distributing scores to students.

Mr. Macon also began a process of discussing common assessment scores with students through one-on-one meetings. This year he was able to reach about half of his class to talk about their performance on the assessment. In some ways, this served as a valuable validity check against the scoring process. Engaging students in dialogue about their exam essays allowed him to compare his own read of their work against their actual thought processes. While essays offered him insight into “what students are thinking, *how* they’re reading the assessment, *how* they’re looking at the rubric, what they’re understanding about it, what they’re not,” he was able to see whether his observations held true in direct discussion with his students.

Like Mr. Urbina, however, Mr. Macon saw that students needed some practice in the interpretation and internalization of their scores:

*I think, where they are culturally is... like they accept it but they don't question. They don't ask, “How else can I improve?” And it's more like me just telling them, it's not THEM asking questions. Although... I asked them, “Hey, do you have any more questions to ask me?” But they don't. So... it's more like me telling them.*

At present, Mr. Macon is finding that, even when given the opportunity to engage in in-depth discussions about their work on the common assessment, students show some difficulty in being able to express thoughts about their performance and/or the ability to understand how their scores connect with their academic performance. Mr. Macon seemed to suggest that assessment scores are viewed by students as less something earned than something conferred. He identified this as a “cultural” orientation, one that needs to be shifted if students are to use the common assessments as meaningful indications of their own progress and achievements. His approach to

this would be to engage students in more active reflection on the assessment process itself. In thinking constructively about their own exam-based writing, reporting, and analyses, he hopes that, by the time they take their final assessment, students will be more familiar with the expectations of the test and how to write to those expectations.

Central to the use of data is teacher practice in test design, implementation, scoring, and analysis. Building teacher capacity to participate in these processes must include opportunities to pilot items, test drive scoring rubrics, and iteratively improve test content and format. Only in the act of carrying out these activities are teacher participants able to engage with assessments with a sufficient level of detail. In Mr. Urbina's words, with experience in all stages of assessment, teachers are able to determine "what their tasks are really asking students to do," and how they are "asking students to get to that place." Practice and experience in testing routines ensure that assessments and scoring mechanisms are appropriately aligned with pre-determined learning outcomes, that those learning outcomes are in step with classroom instruction, and that assessments elicit the type of student responses necessary to accurately gauge their knowledge, ability, and skills.

### **Part III: Data Use in School Performance Monitoring – Impositions on Teacher Autonomy**

Part I of this chapter focuses on the use of data to inform decisions related to student programming and instructional planning. Findings suggest that, while an organizational orientation toward data use is an essential component of effective data use for these purposes, data use is still reliant on teachers' personal understanding of, and fluency with, school data processes. Part II of this chapter elucidates what teachers' personal engagement with data look like in the context of student assessments implemented at Woodson College Prep. This second

section emphasizes the importance of teachers' authentic interaction with measurement development, scoring, and analysis as a way of truly understanding how assessment results can be interpreted into instructional change.

Part III pulls back again from this intensive view of instructionally-informative data to look at factors influencing the use of school performance data in general. In so doing, we enter into a deeper discussion of why the alignment between individual and collective expectations of data use (as discussed in Part I) in schools, can be so complicated by teachers' "ownership of data" (as discussed in Part II).

Using data at the school level requires both organizational and individual orientations toward data use. There must be a collective recognition of common goals, objectives, and questions pursued through data use routines. This collaborative work is dependent on buy-in from individual teachers, but the exchange between viewing school performance in a standardized way, in many cases, stands in opposition to a sense of teacher autonomy and the need to protect teachers' professional space.

### **Teacher Autonomy: Freedom, Power, and Duty**

Mr. Leighton, a teacher at The Academy, described how his own sense of personal autonomy was central to professional identity as a teacher. He perceived District attempts to "standardize" teacher performance in the name of quality control as a serious threat:

*Education is not the place for autocratic tyrants. But that's what LA Unified does. All over the District that's the kind of principal that they install, and they try to force all kinds of edicts down on the teachers. Most of us became teachers because we like the autonomy we have in the classroom. Like our classroom is our own little world where we get to teach what we want to teach, you know? We all love what we do. When outsiders try to force something down on us, they're*

*radically changing the nature of the profession. And that's why a lot of older teachers don't like the new stuff that's coming down the pike.*

When asked to define “new stuff,” he answered:

*Like trying to force teachers to do things in a particular way. Every teacher has their own style. It's a very individualistic thing. You know? But they're trying to turn it into like... an assembly line thing where every teacher does exactly the same thing the same way.*

*And I understand that there is some value in standardizing SOME things, you know? I think, there should be some standardization on CERTAIN things. But, the more you do that, the less personal it gets. And making it personal is what makes it fun. And if you remove all of that from the profession, you're giving people very little reason to stay in it. You know, cause, God knows they are not paying us well. You know?*

For Mr. Leighton, the District is responsible for overshadowing the teaching profession with rules and regulations he feels restrict teacher autonomy. The desire to standardize practice and “force teachers to do things in a particular way” is a District trajectory that he regards as crushing to a profession founded on the style and strategies brought into the classroom by individual teachers. The perception of “outsiders” dictating Mr. Leighton’s practice would render teaching void of meaning and reward. In this way, adhering to external expectations of performance is in direct conflict with his approach to classroom practice.

On the other hand, Ms. Hanley, another teacher at The Academy, saw this ardent notion of teacher autonomy something of a roadblock to better instruction. She explained her observation of this tension between teachers within The Academy as they began to negotiate collaborative work with their departmental colleagues. She began by contrasting this against her previous experience successfully collaborating with teachers at her former school:

*So we would have this one lesson and then we would all look at it, and I'd be like, “Well I would probably do this,” and like, “Yeah, and I could do that too. I think*

*I might also do this.” And then, where, you know, where you start becoming atoms in motion within... that ball. There’s a ball, and you see the ball, but what it really is a whole bunch of atoms working together to create the surface. And so that’s... I don’t think there’s that point yet. We still have these atoms that are completely separate and have not had the opportunity to really understand true collaborative work. You know?*

*And so it’s very difficult to sort of like... What I hear a lot is like, “Well I’m going to have change my whole lesson plan.” Yeah you might. You know what I mean? But they don’t understand what the benefit [is].... It feels like it may be a control issue, it may be a like, “I’ve worked so hard on this I don’t want to change it,” instead of realizing that... You know what? You may actually come out with something that’s so much EASIER when you work with somebody else.*

For Ms. Hanley, the stronghold on teacher autonomy underlying teacher practice at The Academy is a barrier to collaborative work. There is yet an established culture of mutual input or feedback into curriculum or instruction. She sees her colleagues as “atoms” moving along completely separate trajectories rather than in contribution to an overarching shape. A sense of “control” or personal commitment to an idea or lesson currently trumps the notion of investing additional time and energy into adapting and adjusting to the work of others. The benefits of teacher collaboration and a commitment to a “greater cause” are obstructed by teachers’ personal interests.

A similar sentiment was echoed by several administrators within this study who, particularly in schools’ earliest years of operation, found the need to distinguish teacher autonomy from pilot school autonomy. Ms. Heredia, the principal of Belleworth, summed this up in a quick statement when she discussed her work to develop a whole-school improvement strategy:

*I think that at first [the faculty] thought their [pilot school] autonomies meant they can do whatever they want in their own classrooms and be left alone. And*

*I've had to do a lot of [work focusing on] autonomies for the school, not for the teacher.*

Ms. Heredia notes here that, while teacher-led management of the School and teacher-driven instruction may be valued components of Belleworth, they are not the driving force behind its exercise of school autonomies. She believes that many of her staff have, in previous years, confused the sovereignties allocated to pilot schools as license to do whatever they wanted within the domains of their own classroom. Her work over that past year had been focused on shifting the perspective of her instructional leadership team to thinking about school-wide goals and objectives that were endorsed and supported by all teachers, rather than focus on teachers' complete independence in decision-making.

All of these perspectives on pilot school and teacher autonomy seem to reflect a similar theme. Mainly, in executing the vision and mission of a pilot school, how do teachers and administrators balance notions of autonomy with measures of mutual accountability? How does a pilot school, as a collective of individual faculty members and administrators, think creatively about its approaches to data use in a way that honors the sovereignty of school-based decision-makers and teachers as professional individuals while at the same time agreeing on more standardized measures of performance and progress?

### **Something Borrowed, Something New: Teacher Buy-In, Ownership, and Ego**

A common theme observed among pilot schools implementing new strategies toward evaluation and assessment is the expressed desire to ensure that data use mechanisms are tailored to the specific needs, purposes, and approaches of the school. "Outside" influences are approached with some trepidation, and programs, processes, and procedures are not generally

accepted “out of the box.” This is true even when the ideals and intentions of the data use activities being introduced are well-aligned with a school’s vision and mission.

### **The Academy: Adaptation vs. Fidelity**

As an example, Mr. Cooper, principal of The Academy, discussed his introduction of the Teacher Review Program with Mr. Easton (see Chapter 4), both of who contributed to the development of the original program at their former campus.

*And so, our challenge is, how do we, because we know what [the Teacher Review Program] should look like, we experienced it, you know? And the receiving end as a teacher who is going through it, as someone on the team who is part of it, we’ve gone through several cycles of it. Here now we’re giving, we’re sort of handing it over to a brand new group of people who are like, OK, I get it, but you know, like, do they really understand? Do they have the knowledge to, OK, like, this is where to go with it and the nature of the conversation, and what exactly they should reflect on, and how far this should go with that reflection...*

*So I think we need to help them kind of like, look at the areas of things that they can talk about that are very specific, you know? Best practices, and classroom management, you know, and all the things that they craft what it is, as professionals. I want there to be really in-depth conversations and treat everyone as true professionals. So that’s a work in progress.*

From Mr. Cooper’s perspective, a great deal of time and energy was invested in the development of a teacher evaluation program that is sensitive to the professional needs of teachers and which engages teachers as constructively critical colleagues in rich discussion and a peer-based evaluation of practice. He anticipated some difficulty in ensuring that these core program elements translate into The Academy’s implementation of the evaluation program and wonders if his faculty, the majority of who have not benefitted from a fully-immersive experience in the evaluation cycles, will master an intimate understanding of how to engage in the types of professional discussions serving as hallmarks of the program. He sees the need to

build teacher capacity to approach this system with authenticity and to maintain its original vision, capacity that he and Mr. Easton will need to work to provide.

While Mr. Cooper was focused on ensuring the Teacher Review Program was introduced to The Academy with a certain level of fidelity, Mr. Knowles emphasized the need for adaptation to the context of The Academy:

*Well, that was Mr. Easton... and Mr. Cooper taught together at [their former school]. And at [their former school], they came up with this program. They basically developed it, and we've modified it for our program and we're rolling it out, where we've made changes and we're kind of, making it our own. Last year was kind of rough going because it was, kind of, we were taking over HIS program, but now we kind of like, embraced the program. We're getting a better sense of how it's supposed to work and uh... so, it's really, made it our own.*

Though Mr. Knowles worked very closely with Mr. Cooper in building the new-founded culture of The Academy, and considers himself very much a proponent of the new Teacher Review Program, his comment suggests an objection to the notion of simply importing the evaluation program from their former school. He makes a distinction between the previous year, which he considered “rough going” because of a perceived requirement to adopt the outside program, almost as if the act of doing so was an intrusion on The Academy’s territory. In the current year, however, the program had been modified to “fit” The Academy by faculty who felt they needed to make their own changes, so the program was more warmly “embraced.” The key to ensuring teacher buy-in at The Academy, Mr. Knowles seems to state, was to make the program “their own” rather than to reproduce the original. There may indeed have been substantive reasons for this and ways in which the original program did not meet the needs of The Academy’s faculty. However, this perceived need for modification seems to stand in contrast to Mr. Cooper’s emphasis on implementing the ideals and central tenets of the program

with fidelity. Mr. Knowles's comments also suggest that intentions to save time and effort by using previously-developed performance review materials was counteracted by teachers' desire to spend time vetting and adapting those materials.

### **Woodson College Preparatory School: The Expense of "Ownership"**

Dr. Baher at Woodson College Prep acknowledged the centrality of teacher "ownership," not just over data use processes, but over the very data that are collected for purposes of student and school evaluation. She reflected on the development of Woodson's common assessments, particularly for the lower grade levels:

*So all of this investment in the IRLs [Independent Reading Level Assessments]. This is OUR assessment. We're going to the mat for this one. That was SO IMPORTANT [emphasizing with a whisper]. That was the thing that, if I had to go back, I'd say, 100% do that again. Spend ALL that time and energy worrying about what assessments, because... that GROUNDED people's sense of ownership over the measures that would be used to gauge their progress....*

*What's HAPPENED to that measure, which has been really interesting is that particular teachers have taken... well one particular teacher [who teaches Grades 4 and 5]... for two summers in a row has kind of worked with me in the summer as a summer research fellow to clean and own, and set up protocols for collecting and analyzing the IRL data. Okay? And he is really... He's a GREAT example of someone, like he didn't know how to create an Excel spreadsheet, and now he is like running cross tabs, right?*

*Like he's got this ownership, and this FACILITY with data. And he helps... He's the translator for the TEACHERS about... How you input it, if you input as a number, then you have to recode it into this... or a letter, if it's an A through Z scale, then you have to recode it as a number to calculate change... And so he's got all these... great ways of talking about how to collect it, and use it, and analyze it.*

The value that Dr. Baher sees in Woodson's investment in the development of its own lower grade reading assessment is tied to teachers' "grounded sense of ownership over the

measures that would be used to gauge their progress.” By taking part in the development of the instrument, Woodson’s lower grade teachers were compelled to endorse the validity of the instrument, and also found meaning in the ways in which it would be used as a measure of their own progress in reading instruction. As part of this, one teacher in particular found himself inclined to substantially enhance his own technical capacity in data analysis. For Dr. Baher, this highlights a sincere willingness to understand the data and work with it in a way that conveys meaning to instructional practice. Teacher investment in the reading assessment has been key to its maintenance, sustainability, and continual development, as well as in the ways in which the resulting data have been interpreted and incorporated into instruction.<sup>7</sup>

While the upper school departments at Woodson were given similar freedom to select assessments they deemed appropriate measures of student content knowledge, some teachers struggled with the notion of adopting ready-made instruments, even if they were created by the Research Center housed under Woodson’s university partner. Mr. Macon, for example, recounted how the Science Department felt the need to take the Research Center’s science assessment and “break it down,” developing a “new and revised version” that was “tailored” for the Department. Ms. Gilman, from the Social Studies Department, felt that the Research Center’s assessments were a poor fit for her department’s content, and that its scoring rubrics “suck, at least for what we’re trying to do.” Several departments relied on the Research Center for help

---

<sup>7</sup> While the selection of the IRL assessment was at the discretion of Woodson’s lower school teachers, the assessment itself was developed by Fountas and Pinnell (1996). Understanding how to best implement the assessments took some time, but Woodson’s lower school teachers were committed to, and found great value in, doing so (Quartz, Kawasaki, Sotelo, & Merino, 2014).

with scoring their common assessments, but many teachers expressed some disagreement with, or lack of clarity in, how those scores were derived.

Ms. Figueroa, principal of Woodson, detailed the misgivings expressed by some of Woodson's teachers about the ability of the Research Center to adequately prepare and score assessments, using comments from the Social Studies Department as an example:

*Like who needs [the Research Center]? What do those people know? And I'm like well A LOT actually, 'cause that's what they STUDY. Like you might have this sense of, like, this is exactly why, whatever... But, in terms of VALIDITY, or whether an instrument is really a best way to measure something, you may not have enough of a background in that. That's okay, that's not your job. People actually study that and they DO know.*

She pointed out that, while some teachers might not have felt that the ready-made assessments linked closely enough with their own instructional content, this was not necessarily reason to dismiss the contributions of the Research Center for ensuring the validity and reliability of the measurements themselves. From her perspective, some teachers were all too ready to assert their own professional opinions above and beyond the technical knowledge and expertise of trained psychometricians. The notion of teacher autonomy in the development of the common assessments was so strong within the Social Studies Department that it opted not to develop a department-wide assessment. Rather, each teacher within the Department was tasked with creating his/her own instrument although they were all meant to tie into the same scoring rubric.

For the school year studied, however, Ms. Figueroa noted how the Social Studies Department was coming to understand how this emphasis on teacher autonomy actually presents a challenge in understanding how the Department is progressing as a collective:

*I think in the END [the teachers] saw the need for more alignment across their practice, for a better common assessment, because they graded their own.... Because they did it themselves they realized there are challenges in NOT creating*

*a common assessment really. They let everybody create their own. And what happens when people don't follow the agreements.*

*... I think [the Social Studies Department], like, reflected on this. They've realized it, but then they kind of like want to always... which makes sense... cognitive dissonance. Sort of like, rationalize them away. Like, "Well, it's because of this," or, like, "We don't want to lose still our own teacher... autonomy in terms of the assessment."*

*But I think more of them are realizing, but if it's too different then we can't really compare. And this is why data year-to-year is like here, here, here, here, here [uses her finger to trace a jagged, mountain-like shape in the air]. And then, in the end, we don't know students really got better at ANALYZING primary sources, at BUILDING a thesis around the... to answer a particular question. And so I think, like, I'm glad that they're now seeing it like that. Like, is our measure really showing us whether or not students got better at this?*

Ms. Figueroa, working closely with the Social Studies Department in implementing and analyzing the common assessments, sees these particular teachers wrestling through the challenge of maintaining individual control over assessment content and being able to develop a common measure of progress. While teachers recognize that their lack of assessment standardization across the Department produces erratic, incomparable progress data from year-to-year, they are still inclined to “rationalize” these differences in light of the need to acknowledge the specific learning objectives set for their own students. In the end, Ms. Figueroa pointed out, the prioritization of teacher autonomy over common assessment content has resulted in a lack of understanding as to whether and how students are meeting the Department’s learning outcomes, and that the ultimate inability to use these data, perhaps, serves no one well.

Echoing issues raised by Dr. Baher, Ms. Figueroa honored the importance of establishing a sense of teacher ownership over their own assessments and progress measures. However her experiences with the Social Studies Department at Woodson led her to make a distinction

between “ownership” and “buy-in.” Ms. Figueroa viewed the latter as the cultivation of a meaningful understanding of why data use matters in the first place:

*You know, I think that process [of assessment development] is so important, right? And I feel that, people have to own their data. But even before owning it they have to... sort of understand why it matters, or why it SHOULD matter.*

*And that's why using the common assessment, the creation of that is so important, because you have a sense of ownership. But I feel with that also comes a sense of like, propriety over it and almost like this ego that's get built around it. So that if it's not the best thing, then you're like, “Wah! Then it's not... then I don't want anything else, because I didn't create it.”*

*And so I think it's almost how you build the capacity... to understand why it matters, and what should be able to measure how we want to achieve something, and demonstrate that. And then also be open to a variety of ways in which we can do that – some of which can be our own that we create, as long as we can also see the shortcomings in that. Like oh, this measures this, but it DIDN'T give us the WHOLE thing, but THIS other tool will. Yeah, but I think that's hard.*

Ms. Figueroa emphasizes here the tendency of some teachers to interpret “ownership” as “proprietary ownership” wherein oversight of data collected, as well as a sense of responsibility to that data, is confused with complete jurisdiction over assessment content. She also recognizes the limitations of any single measure. Instead, she insists on the need to “be open” to a variety of ways in which identified outcomes might be assessed. These approaches may include both self-created and ready-made instruments, each of which need to be evaluated for their strengths and weaknesses, and which may sometimes require modification and adaptation. But what is essential to keep in mind, argued Ms. Figueroa, is the purpose for which the data are being collected and used. This purpose should drive the determination of tool selection and instrument development rather than “ego,” to which she attributes the inclination to renounce whole tools that are not considered “the best thing” and are not self-developed. She senses, however, that this

sense of “ego” too often supersedes the underlying intention to collectively demonstrate achievement toward Woodson’s goals.

### **Belleworth School of Arts and Technology: Enforcing Standards of Success**

The development of Belleworth’s data use culture provides another example of how establishing school-wide standards has had large implications on teachers and their practice. This year, Ms. Heredia and the ILT decided they would work with Belleworth’s faculty to ensure that learning objectives are developed and physically posted in all classrooms for all lessons. She explained:

*I chose learning objectives because I felt that it was something small to do that has real implications for instruction. Because unless you have learning objectives for every day, you're only going to have them for the end of the semester. And then what's going to happen when you get to Week 15 and 80% of your kids are failing? You won't know what happened because you don't have enough measures. You need daily measures to track learning progress. So, if you have your learning objectives and you connect them to instruction, this is really going to alter the way you think about instruction.*

Ms. Heredia describes her focus on standardizing this relatively small strategy across classrooms because she believes it has larger implications on teachers’ capacity to work through outcomes-based instruction. The requirement to post learning outcomes for every lesson, she argues, is not only beneficial to students, some of who have offered positive feedback about the transparency of lesson “takeaways,” but they also facilitate explicit instructional connections to the new Common Core standards. Further still, Ms. Heredia believes that lesson-specific objectives present ample opportunity to track classroom progress on a finer grain scale than semester-based learning objectives which offers student progress data too late for mid-course improvement.

Despite the perceived benefits of posting daily learning objectives within each classroom, Ms. Heredia had been receiving pushback from some of Belleworth's faculty. On a practical level, Ms. Heredia recounted, "They say it's too much work." From an organizational perspective, she understood that the determination of how much variability should be allowed across classrooms entails careful consideration and must be a collective decision as to what should be the expected "standard of performance." Ms. Heredia herself leans more toward uniformity between classrooms (ensuring, for example, that all classrooms have lesson-based learning objectives posted at the front of the room), because this would show a "minimum performance expectation." At the same time, she recognized that some teachers might see this only as an issue of "compliance" and perhaps an intrusion on teaching individuality and autonomy. She viewed this as a complication for her staff, who had voiced interest in wanting to maintain expected standards of performance as part of their school culture, but who, at the same time didn't believe they should necessarily hold teachers accountable for posting their learning objectives.

Indeed, Mr. Neal, a teacher at Belleworth, expressed his frustration over the learning objectives requirement – what he viewed to be a simple measure of accountability and one that he did not fully understand:

*I think [faculty] were more willing to do things whether or not they agree with it or NOT just to save FACE because we didn't know how they're going to be evaluated by it. Yeah, so documentation of THIS, and documentation of that, it became more of... let me make this paper trail, let me do this so... it can be seen that I'm doing this and I'm doing that.*

*Like, one of the biggest complaints... was about having the agenda on the board, when I would have it in my [PowerPoint] slides. So I wouldn't have [it on] the board there, and we'd be looking at that anyway. This is what we're going*

*through, this is what we're studying, as has already been explained, but it's just like...*

*And I know there's certain things that we have to do. And we kept being told we don't have to be cookie-cutter, you know, but we have to do these, and we have to do those things. So when I DID finally put an agenda up... it was kind of like... "Well, that's not good enough." But I went with our Common Core standards that we have, that we're gonna' use for the class. So I would pick one of those standards to incorporate it into what I was teaching, and I put it on the board along with the California State Standards that I was teaching, and it was still deemed not sufficient enough. Like that's what we're DOING, that's what we're going over.*

Mr. Neal saw Belleworth's focus on learning objectives primarily as a measure by which to evaluate teacher performance. From his perspective, the need to physically post learning objectives in every classroom, and faculty adherence to the new policy, was evidence of a new dynamic emphasizing documentation and the creation of paper trails to ensure teacher accountability. Although Mr. Neal believed he was in compliance with this new expectation, i.e., by posting his agenda, utilizing Common Core Standards, and having California State Standards in his classroom presentation slides, he had been told that these efforts were "not good enough." Mr. Neal interpreted this as a consequence of his deviation from a "cookie-cutter" classroom presentation of learning objectives (i.e., presenting his agenda as a PowerPoint rather than physically posting it on his classroom walls), as well as spurious specifications as to what needed to be posted (i.e., the standards he has posted do connect with the lesson, and he had already verbally explained to his students what they would be doing).

What seems to be missing from Mr. Neal's interpretation of Belleworth's focus on learning objectives is what Ms. Figueroa identified as an understanding of the purpose of this particular teacher performance measure. Mr. Neal finds himself "going through the motions" in compliance with the new approach, but without a comprehensive understanding of why this has

become a minimum expectation of performance. As such, he seems to have missed the underlying strategy of stimulating teachers to develop lesson-based outcomes on which to gauge student achievement. Instead, he has only connected his lessons to a more general set of standards.

In contrast, Mr. Nuñez had his own misgivings about Belleworth's new minimum performance expectation, although understood its intent and the overarching need to move in the direction of articulating his lesson objectives. His deliberation was evident in the discussion of an important drawback of the new policy:

*You know for math, some of my favorite lessons are where [the students] just come in and... they have no idea what it is, but they have one problem on the board, or have one activity on the table, and then it's more like an exploration. You know? And the question is, okay, so what do [you] think you guys are going to learn? What do you think this model is trying to teach you? Or what is this question trying to get to you?*

*So... that new policy that we've adopted kind of exes out that whole exploration [by] telling them what they're going to learn (laughing)! So then I'm scratching my head and saying, wait a minute so... that great exploration activity on say, the area of a parallelogram, or the area of shapes, that's out the door. Because now what am I going to do? How are they going to... How is that going to tap into their curiosity, and their imagination?*

*But again, you know, there's pros and cons and... if it's something that we voted on, then regardless of whether I voted yes or no, I mean, we have to do it, you know? Just another thing we have to do. The only good thing is that we're not, we're not TRYING to change the way people are teaching, but you know, with the implication of the new [Common Core] STANDARDS, we HAVE to. You know we have to. We have to KIND of modify the way we used to teach. So...*

*And it's not much on the whole pilot school, like the... or how we want curriculum to change, it's just... new standards, so we have to change.*

Mr. Nuñez raises an interesting unintentional consequence of demanding uniformity across classrooms with respect to posted learning objectives. If all teachers are expected to

clearly delineate their learning objectives at the start of each lesson, how might he also accommodate one of his favorite pedagogical strategies of employing student-directed investigation, questioning, and reasoning in the exploration of an undisclosed learning objective? Mr. Nuñez feels that one of his best classroom activities has been rendered unusable. In our discussion, Mr. Nuñez did not ruminate on this drawback, but rather focused on the more general “pros and cons” of the policy. He recognized that, while he may have his own personal issues with the strict implementation of lesson objectives, the intent of the policy is not to modify his approach to instruction. Rather, he saw the demands of the new Common Core Standards as the impetus for change, as well as the resulting need for all teachers to subsequently shift their pedagogical approaches. His own “vote” on the Belleworth’s learning objective policy came second to decisions collectively agreed upon by the faculty, and to the school’s response to new standards.

While Mr. Nuñez may have taken issue with some of the ramifications of Belleworth’s new learning objective policy on his own practice, Ms. Heredia feels that other teachers refute the idea simply because they view it as an infringement on their individuality. This hard-lined position on teacher autonomy is, in her perspective, somewhat misplaced as Belleworth attempts to establish a culture of school-wide accountability. She noted:

*Sometimes I feel like [the faculty] think they're defending their autonomy in some way, their classroom autonomy, and their individuality as teachers. But I feel like it's in the wrong space. Like, you don't defend your individuality there [in refusing to post learning objectives]. You defend it in the projects you have kids engaging [in], the type of content that you choose to present to the kids – you know what I mean? And your style, your strategies[that] you use, but not on things that should be formal things in every classroom, like basics.*

For Ms. Heredia, the posting of learning objectives in every classroom is, in some ways, only one physical element of the classroom environment. She wants to ensure that every teacher makes lesson learning objectives clear to their students, but expects that this will impact teacher planning more than their own pedagogical style or curricular approach. The stubborn objection to uniformly posting learning objectives at the front of each classroom thus seems to her a misplaced demand for autonomy.

### **Public Accountability**

While there was certainly an element of practitioner perspective at play in Belleworth's ongoing debate over whether and how learning objectives should be displayed in all classrooms, Ms. Heredia raised an important point in terms of negotiating a school-wide approach to data use and accountability. Namely, as a school begins to construct the performance outcomes to which it both aspires and will hold itself accountable, where is the appropriate juncture for teachers to forfeit some of their classroom autonomy for a common cause? If each new outcome-based strategy has larger implications on teacher practice, where should a school start? Which "battle" is worth fighting for? These questions seem to precede a collective understanding as to the purpose and intent of school-wide data collection and performance monitoring.

Ms. Heredia, for example, discussed the challenge of getting Belleworth's faculty to agree that, for at least the day of their District-led Pilot School Review, all teachers should make sure their lesson-specific learning objectives are posted somewhere in the classroom:

*Because I feel like we... at least for that day, we should have a structured way of writing learning objectives so... "We [faculty] don't see why we should fake it. Why are we going to fake that? Why are we going to fake the learning objectives for somebody else?" And I'm like, "Well it's not like FAKING it, it's just like...*

*we're saying that we're working on this – it should be... we should be able to show it.”*

*.... So we were having the conversation and they're like, “Well everyone should have them. But the problem is everybody DOESN'T.” So then, what are the minimum things that we're going to say people need to have? You know? And then people were like, “Well I don't know if we should... say that because that's not the most important thing in the classroom.” So then, it becomes okay, but if we've been focusing on that all year, and you can't see it when you walk in the classroom spaces, what is that going to say about our work?*

Although Belleworth's faculty decided to make learning objectives a focus for the year, by the time of its Pilot School Review in the spring semester, they were still unclear as to whether they should require all teachers to have these posted in their classrooms. Ms. Heredia recounted the sentiment expressed by teachers who felt that posting learning objectives for the Pilot School Review would be a dishonest representation of their actual practice; to do so on review day would be “faking it” rather than evidencing capacity and capability. She watched as the conversation about learning objectives devolved into a discussion about which key elements – i.e., which minimum standards – should be evident in all classrooms. Ms. Heredia found herself reminding staff that they'd been focusing on learning objectives all year, and that faculty should be able to hold themselves responsible to their self-determined goals. The propensity for staff to construct new standards rather than adhere to their original plans would undermine “their work.”

Even following the conduct of the Pilot School Review, several teachers questioned the validity of its findings in ways that impeded the integration of results into instructional improvement. Three teachers wondered whether the limited classroom observation periods (occurring either at the beginning or the end of a class period), and the unequal rotation of observation teams among the classrooms (such that not all observation teams were able to

observe all classes) limited the accuracy and reliability of the data collected. At least one teacher described his misgivings of some of the evaluation criteria which he felt were not adequately deconstructed among faculty in the analysis of the Pilot School Review results. Irrespective of the perceived validity of findings, however, some teachers still endorsed their use. In step with previous research suggesting that evaluation consumers tend to accept those results that reinforce their own beliefs (Weiss, 1995), several of Belleworth's ILT members seemed more inclined to support Pilot School Review findings that reinforced their own previously-held views of Belleworth's progress. While Ms. Salceda conceded to some indicated areas of improvement because of "observations I've made from my own classroom" for example, Ms. Nava found herself backing the Review's criteria because, "For *me* I think these are things that I'm working on. Like, I have my goals for the year that *I* want to work on. So when the Pilot School people came in, it was kind of like, oh okay, so I'm hitting the right goals." In these instances, the exercise of personal judgment appears to outweigh more collective concerns over data credibility.

Like Ms. Heredia, Ms. Figueroa also found many teachers at Woodson College Prep reluctant to publicly announce their common assessment findings or to hold themselves accountable to independently-developed benchmarks or patterns in year-to-year data. For these teachers, the common assessments served as an important source of self-reflection, but were not pieces of information that are comfortably published as a formal measure of achievement. In some cases, teachers were hesitant to share common assessment results with colleagues outside their own departments. Ms. Figueroa was frustrated by the aversion of some teachers to analyze their common assessments in this way, which she believes is the primary purpose of the instruments:

*Because the idea is to show the strategy, show [data] use, again... what we're missing. I'm like, aren't we presenting what students learned? You know, like, why aren't we being more open about that? Like, "AND we saw a 30% increase," or "It didn't work! Kids stayed the SAME, and yet, we saw THIS. It didn't translate into this, but it definitely, we could see THIS."*

*There are still some departments that are like, "Oh, we don't have that at all." And I'm like, then what's the POINT? It's not just... I mean seriously, the point is not just could you just reflect for yourself on your [data]... No. The point is, what do students get out of all of this work that you did? With them, for them. Right?*

*It's still about, like, "Well my reflection and what I'm learning." I mean, that's important, and I understand that, of course. But if I'm not like making anything else where I can SHOW this, then to me it's like, what's the point? You know like, in the end use it should translate into graduation, and kids reading, higher levels of bilingualism, bilingual literacy. How are we measuring that? How do we show that?*

When asked if there are just some people for whom those measures don't capture effectiveness, she explained:

*But see, those are our own measures. Like, I understand if you don't love the [standardized English language development test] because that's not an instrument you created. But the ones WE'RE saying we love? No, we better care. We better know that half of our kids, again, didn't meet their benchmark. And that should tell you something. And it should push you to action. That's MY thing. Okay fine. You don't love whatever. But... we have to do it, the STATE [test]. But I'M talking about the ones WE give. That we say are like SO amazing, and give us SO much information. Like what we do with it? You know what I mean?*

*And you know we didn't create [the common assessments], but [they are] things that we value because they definitely inform our instruction. And they SO have informed your instruction. Show me then how we're using them in a way that is really helping us make better choices.*

Ms. Figueroa here emphasizes the importance of analyzing student progress through data and using assessment findings to "show" changes in student achievement. She believes this is the kind of empirical picture teachers need to really gauge what impact the work they have

committed with, and for their students has affected. She is careful to contextualize student achievement results, however, and underscores the importance of what can be learned even if benchmarks aren't met or upward trends aren't observed. Students may have moved or improved in ways not detected by the assessment, but the assessment serves as an important baseline of discussion. Using the common assessments as a tool for self-reflection on teacher practice bears value but misses the ultimate purpose of holding teachers to their goals and objectives in ways that are observable and communicable. Such data should be used for more than just internal instructional adjustments but should also be used as evidence of student progress.

In response to my suggestion that some teachers may not feel that standard, aggregate measures of progress adequately capture “effectiveness” in teaching and learning (certainly, some participants within this study voiced their skepticism of “numbers”), Ms. Figueroa made the distinction that Woodson’s teachers have gone through the process of carefully selecting their own measures of progress. If anything, she argued, these should be the standards to which Woodson faculty hold themselves accountable. While other measures may be doubted for their validity, or their applicability to practice, Woodson’s common assessments – selected, administered, and, now scored, by faculty – should be collectively considered a valid measure of student achievement.

In reflection, Ms. Figueroa understood the reticence to use data as a measure of performance as an attachment to ego. To hold oneself “accountable” also means to be receptive to constructive criticism and open wide to self-reflection, which is, inherently, intimate feedback. Thinking about the common assessments, she observed Woodson’s teachers progressively coming to terms with a variety of outcomes. She emphasized, however, that alongside “being

okay with where you are,” teachers must still “still focus on where you need to BE.” She commented on this thought process:

*And I think that that's really hard because teaching is such a personal thing. You're putting like your heart and soul into it. Then I think what I found... is that it's so hard to be like, "But I put all my heart and soul into it and you're still saying that students aren't doing what I wanted them to do?"*

Ms. Figueroa understands the profession of teaching as an extremely personal undertaking and one that is naturally susceptible to teachers' instinctive reactions to evaluative findings. Regrettably, the extensive work invested by a teacher in his or her classroom does not always translate into improvements in student achievement. Some cognitive dissonance results when measures of student progress do not reflect such intensive investments. But, rather than speculate on the validity of the measures themselves, Ms. Figueroa suggests that such opportunities provide critical moments of learning:

*But then look at, so what are we LEARNING from that so we can then be who we're meant to be? You know, like, that's when you really see what you're made of. Not in the moments of success. But in those moments of, like, (sighs) you know it's all falling apart, so then WHAT, you know?*

*I think that obviously, because there's ego, because there's this pride around and hubris, that's why it's so HARD to be like, okay. Well it's pride, all right, done. Now let's think about, so how do we get up again?*

While Ms. Figueroa sympathizes with the intrinsic defensiveness resulting from poor evaluation or outcome results, she also attributes such reactions to a sense of pride and an ego surrounding the work of educators. It is the educator's responsibility, from her perspective, to translate critical findings into constructive improvements in teaching and learning. The “ego” is something that she sees as standing in the way of this conversion and what obscures the utility of data that do not immediately reflect the success of strategies, interventions, and innovative

approaches. There is something to be learned from these moments, she argues, and they provide important opportunities to exhibit resilience, persistence, and a commitment to valued goals and objectives. Instead, what Ms. Figueroa experienced was a tendency for many teachers to disregard student outcome data and a hesitancy to evidence progress with student outcome data:

*I think making the connections that everything that we're doing in the end should be VERY connected to student outcomes in terms of achievement. Sometimes people really shy away from that. I mean, they're willing to grapple with all of these issues, and, like, "But let's try to improve this." But then once you say, "Okay, let's see if it WORKED as per grades, as per passage rates, as per this assessment," then they're like, "Why? Why would we want to look at that?"*

Ms. Figueroa noted that, while the intention of faculty to improve student achievement and progress is certainly present, a commitment to measurable outcomes is not. For Ms. Figueroa, such outcomes serve as proof of concept, a measure of a theory of change. But for many teachers at Woodson, she observed a distrust of student outcomes as a defense of ego. Importantly, she viewed this discrepancy as a stage Woodson could work through. Woodson's strength in identity, mission, and motivation would drive teachers toward needed improvements:

*We're going to DO it. I mean I do have that belief too, that we are going to do it. Because of the people we HAVE. I'm like, we are going to do it because we are who we ARE. You know, like this identity. And that identity is important to maintain.*

*But not without remembering that we're also vulnerable to like, like, I don't want us to get into this ego trap. You know? There's a certain humility that we need to approach the work with, too.*

Ms. Figueroa sees Woodson's faculty as a dedicated group of professionals with strong ideals and backed by a strong sense of identity. She believes these features are key characteristics of the School and will both compel and propel teachers through the work of improving student achievement. She emphasizes the importance of balancing both identity and ego, however, in

being able to constructively reflect upon measures of progress. She believes that a certain sense of “humility” is also essential in the process of deciphering which approaches have been successful and which less so.

### **A Parallel Universe: District-Level Oversight and School-Level Discretion**

Parallel to the discussion of teacher autonomy and school-wide accountability is the consideration of pilot school autonomy and District-wide accountability. District perspectives on school performance monitoring add yet another layer of complexity in understanding how data are used and regarded by schools. A special unit delineates that District managers within Superintendent’s ISIC are responsible for guiding the establishment, development, and management of pilot schools. Part of its function is to ensure that pilot schools are operating in accordance with the terms and agreements of their memoranda of understanding in the conduct of a formal Pilot School Review, a process that involves school site visits, classroom observations, and the evaluation of teacher and school performance against multiple standards by various school stakeholders. In fulfilling this role, the new Director of Autonomy and Accountability, Ms. Macia, is thoughtful in her approach to cultivating stakeholder voice and buy-in around the Pilot School Review. She is sensitive to the notion that, in order to meaningfully engage individual Pilot schools in the evaluation of their performance and to encourage their use of Pilot School Results to inform further improvements, a certain degree of adaptability is required on the part of the District.

Having been both a pilot school principal and an instructional director, Ms. Macia was well aware of the differences in the ways in which the Pilot School Reviews are introduced to each school. At the outset of the 2013-14 academic year, she reflected on the implementation of

the Reviews, commenting on the role of the instructional directors to guide classroom observations and reach consensus among school stakeholders about the final findings to be reported:

*I think [the reviews] are done differently [from school to school].... There are some directors who believe that there need to be more external members on the team, and probably don't spend as much time, um... coaching team members in how to gather unbiased evidence. And maybe take more of a traditional approach in the debriefing of the conversations.*

*In general, I think that what you might find is a difference between facilitation styles that might generate more voice... or less voice. They're subtle differences, but differences that might reflect your philosophy about how to manage and how to facilitate conversations, and for what purpose.*

When asked who was responsible for facilitating those discussions of consensus in the reviews that she was part of, Ms. Macia replied:

*In the reviews that I was part of, it depended. So, in my network of eight schools, there were some schools where the facilitator... where the principal was very comfortable in facilitating and I thought had a mindset conducive to nurturing and supporting the pilot school philosophy. One of them being distributed leadership... democratic practice, and by that I mean including the voice of students, parents, and teachers, those closest working to the students. In other schools, there were perhaps newer principals or principals not **as** comfortable, and so I modeled some of that for them.*

From her experience, Ms. Macia anticipated differences in the approach of instructional directors and principals to the Pilot School Review. While the procedures for each review might appear similar from school-to-school, subtle differences in the way school leaders manage and facilitate conversations among school stakeholders, she noted, seemed to be a reflection of their varying philosophical standpoints on how these dialogues should be carried out and for what purpose. The issue of “voice,” i.e., who represents the school body, how they are heard, and the ways in which their perspectives are reflected in the summation of a school’s performance, was a

central point of distinction for Ms. Macia. In some cases, she found that principals were more proficient at integrating stakeholder voice into the Pilot School Reviews and, in other cases, that modeling this type of dialogue was helpful.

Ms. Macia explained the fairly intensive processes of preparation previously undertaken with schools for which she was the Instructional Director. This involved careful conversations around observer bias and how to “objectively” script teacher and student activities observed in the classroom. She coached participants to save their evaluation and analysis until groups could collectively discuss how their collections of evidence might be organized in the context of the performance rubric. As the Director of Accountability and Autonomy, Ms. Macia built some facets of stakeholder voice into all Pilot School Reviews by conducting focus groups with students, parents, and teachers. However, she simultaneously recognized that there are limitations in the extent to which her own philosophical approach can be standardized across reviews. She commented on whether the review process conducted this year at Belleworth was typical of what she saw at other schools:

*I think that, there were probably more people there than... on average. A few more people. But in terms of the process... um... yes. With the exception that, usually, the Instructional Director takes a little bit more of a lead in explaining the process. And I was going to jump in, but I was being deferent to the Instructional Director's position and authority.*

Ms. Macia observed that the Instructional Director for Belleworth took less of a lead in explaining and framing the Pilot School Review process than she has typically observed. Issues concerning reviewer perspective or scripting and debriefing guidelines were not explicitly discussed as Ms. Macia might herself have done as an instructional director. However, Ms. Macia made a conscious decision to defer to Belleworth's Instructional Director out of respect

for her position, authority, and relationship with Belleworth's administration and faculty. These are important political elements of the Pilot School Review process to acknowledge, even if they mitigate "consistency" across reviews.

In addition to her political observation of title and "authority," Ms. Macia also recognized the need to maintain a certain level of flexibility in the review process:

*To your point about the actual observations... some would have very extensive amounts of time in the class and others less. So, I didn't actually get to that point when I was trying to... create guidelines for more consistency but... I wanted to make sure that there were at least some foundational pieces that were common.... Largely, discretion is given to the Director BECAUSE there's a danger in making everything standard. And that is that; you may not address the needs of the school. So, in trying to find that balance, we find that things WILL be different, that the rules WILL be carried out differently. And from my perspective, that's OK.*

For Ms. Macia, the discretion of the Instructional Director is an imperative component of the review process. To standardize every element of the review process would detract from the Director's ability to flexibly respond to each school's individual needs. In considering the balance between codifying the entire Pilot School Review process (a way of reinforcing the standardization of findings between schools) and maintaining a certain degree of site-level flexibility, Ms. Macia accepts that a focus on only the foundational components of the review process will naturally give way to differentiation in implementation across schools. From her perspective, this is a necessary tradeoff.

Indeed, her acknowledgement of school-specific needs is deeply rooted in the notion of stakeholder buy-in to the review process itself. In discussing the differences between structured and semi-structured performance rubrics, she noted:

*You know, you're going to gain something with having one approach and you're going to lose something. So what we're losing is consistency across. Something*

*we GAINED is... kind of a mindset that really this is about INTERNAL accountability. That was the message that I wanted [schools] to come away with.*

*Because external team members can comment, and we can make assessments, but if the school's leadership team doesn't take OWNERSHIP of it, then... I don't see the purpose in it.... So why make people FEEL like, oh no, we've got to live up to this... you know, we've got to FIT what we're doing into this... rubric that somebody else created?*

Ms. Macia emphasized the importance of cultivating a sense of “internal accountability” amongst schools. In her view, this requires a feeling of ownership by school stakeholders – a genuine regard for scoring criteria that the school’s leadership feels is valid and relevant to the school’s vision of teaching and learning. She recognized that the imposition of standards that are perceived as external to these values may be acknowledged out of compliance – external team members will comment and assessments will be completed – but they will ultimately lack meaningfulness for the school in its own quality improvement. Ownership over performance data, therefore, relies on a direct translation between a school’s own activities and the criteria on which they are evaluated. Ms. Macia intentionally integrated room for a reflexive, responsive approach to the reviews as way of cultivating stakeholder buy-in to what she hoped to be a constructive process of feedback. The ways in which pilot school performance is assessed must yield to the variation in approaches to teaching and learning expressed by each school’s unique vision, mission, and teaching and learning objectives. Without this flexibility, Ms. Macia fears that the Pilot School Review data will be regarded as irrelevant, invalid, and ultimately, useless.

Implementation of the Pilot School Review with this degree of introspection, however, is often complicated by practical limitations. Ms. Macia recognized that the ability to successfully carry out an observation-based review that is meaningful to schools requires technical expertise on the part of school leadership. Principals must be able to facilitate the objective collection of

evidence, calibrate scoring and language amongst observers, and translate observed practices into performance standards. But principals, she pointed out, are “extremely BOMBARDED” and “SO overwhelmed” with school-based work requiring “deeper and greater leadership,” particularly in pilot schools where human resources are extremely limited. Ultimately, Ms. Macia found that principals defer to the guidance of their instructional directors or to the Central Office. She noted, “[Principals are] not interested if it’s a renewal, if it’s a review, a one year, three years, five year... They just say, we need to get through this, let me just get through this. OK? And we’ll do it to the best of our abilities.” The time, resources, and expertise required to implement an ideal system of review with fidelity is more than what most schools can afford.

### **Cross-Case Insights**

While much of this study has been dedicated to the illustration of how teachers readily make use of various types of data in the course of their own instruction, it is also understood that the standardization of data use processes as a whole-school strategy is a complicated endeavor. While teachers may generally commit to data use in decision-making thus endorsing an overarching approach to data use, whether and how data are used is left to each teachers’ discretion. The affirmation of data use processes includes different stakeholders’ determinations of what data are credible, their respect for decision-making processes, and their agreement as to how data ought to be used and for what purposes. This section extends the discussion of the individual-collective dichotomy in schools by focusing on teachers’ sense of autonomy and mutual accountability with respect to school, teacher, and student performance standards.

The independent production and use of school data to self-monitor and assess performance is reliant on several steps; schools must: articulate their prioritized teaching and

learning outcomes; determine appropriate goal-lines against those outcomes; plan and implement activities, interventions, and strategies that map to their goals and outcomes; and, determine the ways in which they will measure progress toward those outcomes. These stages are fundamental for effective data use wherein relevant data are identified, collected, analyzed, and are expected to be used in conversations around school progress as indicators of achievement and needed improvements. This section has shown how, even when these steps are in place and a school has developed an explicit strategy to respond to and assess collectively-determined goals and objectives, a lack of teacher “buy-in” or a copious sense of “ownership” over the use of school performance data can inhibit its use.

Examples from The Academy, Belleworth School of Arts and Technology, and Woodson College Preparatory have all shown differences in perspectives toward data use between teachers and administrators. The intention is not to emphasize these variations as a divide (certainly, there are teachers and principals who share similar viewpoints), but instead to highlight some of the complexity in respecting both teacher autonomy and a sense of mutual accountability within schools.

As seen within The Academy and within Woodson, teachers’ demand for ownership over data collection tools and processes nearly precludes the adoption of ready-made materials. This stems from the perceived need to adapt measures of student and teacher progress to the unique classroom and school contexts in which they are applied. Teachers’ desire to maintain proprietary rights over data collection tools, however, has also threatened the rigor with which teacher and student performance are measured. In the case of The Academy, its TRP is not being implemented with fidelity and, as a result, is missing components considered fundamental to its encouragement of authentically engaging and constructively critical peer reviews. Some teachers

at Woodson are finding that their insistence on developing their own student assessments has produced test results incomparable from year-to-year and from classroom-to-classroom.

The development of a school-wide strategy for improvement is currently underway at Belleworth. Administration, teacher leaders, and other faculty are still in the process of defining what it means to allow indicators of school performance to both assess and guide teacher practice. Teacher perspectives on whether teachers should post their lesson objectives in their classrooms range in agreement. While some view this to be a valuable activity for both students and teachers, others see the requirement as just another layer of compliance. While some teachers believe the posting of learning objectives should be a school-wide feature, at least one teacher pointed out the potential drawbacks such uniformity may have on his own pedagogical approach.

At both Belleworth and Woodson, it was sometimes difficult to garner a sense of mutual accountability to school-wide standards of performance. Although Belleworth's faculty decided to focus on teachers' articulation of lesson learning objectives, teachers characterized the requirement to post learning objectives for its Pilot School Review as an exercise of "faking" conformity for the District. Some faculty at Woodson appreciated student assessment results as informative for class-based instruction, but declined to share these data with their colleagues as the basis for determining school-wide progress. In both instances, there seemed to be some reticence among teachers to hold one another responsible in either implementing strategies for improvement or in tracking strategies for improvement outside of their own classrooms.

As shown across these three cases, cultivating a sense of school-based accountability to improved student achievement and teacher practice has necessarily entailed acknowledging teacher autonomy. Teachers' engagement and "buy-in" to systems of accountability and

evaluation are strongly linked to their own self-determination in supporting school-based outcomes and objectives, data collection activities, and the interpretation of results into classroom practice. As has also been seen, however, teacher-led data use cycles are complicated by natural limitations in evaluation and assessment technical expertise (see Part I). As has been described in this section, it is sometimes also difficult for schools to objectively accept and reflect upon negative assessment and evaluation results. For all these reasons, teachers have been seen to subsequently challenge the credibility of aggregate data in their accuracy in representation of whole-school performance. While Ms. Figueroa discussed this as an issue of “ego,” Dr. Baher mentioned it as an issue of “trust” or “faith” in the ability of evaluative systems to evidence student growth (see Chapter 9).

But there is also the question of what substantiates “mutual accountability” within a school wherein teachers and administrators within a school work together toward common goals and objectives, collectively holding themselves responsible for their attainment. This cooperative sense of responsibility is reliant on more than a fear of those negative ramifications resulting from non-compliance, but is instead founded on relationships of trust and understanding among teachers, a dynamic less understood outside specific school culture contexts. Indeed, the ability to maintain a strong sense of “accountability” within a school relies very much on to whom stakeholders perceive themselves accountable. Most certainly, many teacher and administrative participants have expressed their commitment, first and foremost, to students, parents, and the community. Teachers have, in several instances, also expressed their deep sense of responsibility to their colleagues. On the other hand, a few teacher participants have mentioned that accountability to their administrators is not of particular concern. Responses to District mandates have been frequently regarded by both administrative and teacher participants as a matter of

obligation rather than of duty. As such, in discussing “accountability” at the school level, we must take care to recognize the many different perspectives influencing its meaning and interpretation.

Indeed, at the District level, discussions with District leadership reveal similar conversations around the promotion of pilot school autonomy while simultaneously acknowledging overarching standards of minimum school performance. The nature of pilot schools is such that each campus is intended to approach teaching and learning through the lens of a unique vision and mission. Each campus is comprised of faculty and staff with varying degrees of experience and capacity. As such, a performance evaluation that does not inherently acknowledge these school-to-school variations is anticipated to be politically untenable. Rather than being viewed as a reliable assessment that is consistently implemented across pilot school campuses, the Pilot School Review, for example, would more likely be considered an imposition of externally derived standards ill-fitted to a school’s unique needs. To promote use of the Pilot School Review findings in school-based decision-making, Ms. Macia explained how the credibility of the evaluation must be endorsed by stakeholders and earn users’ confidence in the data’s relevance, meaningfulness, and application.

However, Ms. Macia also knew that the use of District-collected performance data was likewise dependent on the efforts of principal and teachers to disentangle and translate data into organizational and institutional change. She recognized that the resources available to schools, in the sense of time, funding, and technical capacity, are known to be in short supply. Stakeholders consistently emphasize that such resource constraints are especially pronounced in the context of pilot schools which operate with fewer administrative personnel than conventional public schools. As Ms. Macia point out, pilot school principals are particularly overwhelmed by

increasing demands for “deeper” leadership and technical guidance. As a result, fuller personal engagement in the Pilot School Review process presents considerable challenges to principals and teachers. Nevertheless, richer discussions with school participants about how Pilot School Review data are collected, analyzed, and interpreted would go a long way toward participants’ understanding of their role in collecting data and negotiating scores, assuaging teacher and principal concerns over the reliability and validity of data, as well as in promoting stakeholder use of the data. In the absence of additional time, energy, and the technical capacity to engage in a more in-depth review and analysis of the Pilot School Review process and findings, a summary of the Pilot School Review’s methodology, assumptions, and limitations are, at minimum, warranted.

In terms of how schools might build a culture of mutual accountability, several participants expressed the value of teachers’ ability to push one another in improving their professional practice. For example, in the upcoming school year, The Academy’s Mr. Cooper and Ms. Heredia at Belleworth both looked forward to onboarding new staff who have a strong track record of excellent classroom instruction. They believed these teachers might serve as models for current staff, raising the bar of what “good performance” should look like. Mr. Macon discussed how a culture of mutual accountability was growing within Woodson as an underlying characteristic of the school rather than an explicit expectation:

*Because even though it's not spoken... you know, what kind of responsibility to have as a teacher AND as a member of the school... you get to see at the end of the year. And you get to see the kind of product [e.g., common assessment scores] you have available for the rest of your colleagues.*

*So... If you're personally doing well (laughs) and you're producing for the rest of the members of the community, then you know you did your part, and you give your own self a pat on the back for that. But... again, no one asked you to go this*

*deep into... you know, into the profession. But we all are, in some way expected to get there even though it's not asked.... So, I think it's just an atmosphere here, yeah?*

When noted that this seems to be a culture of and that Woodson set high expectations, Mr. Macon expanded:

*Yeah, yeah we do. We do. (Laughs) We also talk about, you know, when we don't. When we need to mess up, or what we need to improve on when we mess up.*

From Mr. Macon's perspective, there is a tacit understanding of the level of performance expected among Woodson faculty. At the end of the year, when student progress is reviewed (either by way of common assessment results or in reflections built into the PDSA process), there is also an opportunity for teachers to exhibit to their peers what they have accomplished over the year. Mr. Macon characterized this as a positively incentivized experience, wherein one can give oneself a "pat on the back" for having done his/her "part." That there is some perception that each individual has "a part," however, suggests that each teacher within Woodson accepts a degree of responsibility for their contribution to the quality of teaching and learning within the school. Improvements to the profession of teaching are not publicly discussed outcomes, or expectations outlined in a teacher's job description, but are built into the "atmosphere" of Woodson. Beyond recognizing what has been accomplished, as Mr. Macon pointed out, is the ability for teachers to collectively acknowledge when they have "messed up" or in areas of practice that warrant improvement.

Dr. Baher confirmed the existence of this unspoken expectation in her discussion of whether teachers were selected to work at Woodson based, in part, on their inclination to look to data as a measure of student progress:

*I think one thing that teachers notice and talk about when they come to our school is that... you're more accountable. In, like, an authentic sense, right? Like, your... practice is going to be public, people are going to be visiting, you're going to have to go through this evaluation process that is, you know, kinda' up close and personal, but also... It's also going to require you to be a real professional. Like, can't just blow this off. This is a REALLY professional moment that you take seriously – the collection of artifact data about your practice, and have people coming in and dialogue, and observe, and finish the [Instructional Quality Assessment].... And would you join into a staff voluntarily knowing that if you weren't as secure? Maybe not.*

Here Dr. Baher discusses the “authenticity” of accountability at Woodson where teachers are held not just to metrics as measures but to evaluative systems and processes upheld and endorsed by the entire school. From informal classroom visits to intensive teacher evaluations, the structures which effectively monitor and assess the quality of Woodson’s teaching and learning regard the teacher as “a real professional.” To take oneself seriously, then, requires that teachers take those evaluative systems seriously. In addition to the peer-based expectations Mr. Macon identified as characterizing Woodson’s culture of mutual accountability, Dr. Baher pointed out that an implicit level of professionalism demanded by these processes and procedures serves as an important screen for incoming teachers. Combined, these teacher-held and systems-led expectations encourage a culture of mutual accountability at Woodson that has become self-propagating.

Woodson is an example of how mutual accountability is being established, but also how relationships of trust around data are complex and take time to build. Across the three schools, the expectation is that data should be used to objectively gauge and assess school performance and, that by nature, such examinations are intended to be critical. But to be constructively critical – in ways that effect real changes in teaching and learning – the “external” views of student and teacher performance posed by school-level data must be internally accepted. This necessarily

entails mediation between notions of teacher autonomy – the recognition of what teachers can control in their classroom – and school-wide standards of performance – a mutual understanding of what teachers should hold one another accountable to. The experiences of The Academy, Belleworth, and Woodson in institutionalizing faculty reflection on performance data provide important insight as to how schools in different stages of development are navigating through this space.

## **CHAPTER 8**

### **The Strength of the Anecdotal: Professional Judgment as “Second Tier” Evidence**

#### **Introduction**

Schools, principals, and teachers are increasingly encouraged to turn to school-based data as an essential point of reference for decision-making. For some advocates of the data-based approach, the use of empirical data is a far more consistent, reliable, valid, and objective informant to judgment than the traditional reliance upon subjective, untestable strategies based on instinct, intuition, or educational trends. Indeed, all of the schools in this study recognized the value of systematically-collected data for purposes of tracking and monitoring school performance, determining the potential impact of school-based interventions, and evaluating the effectiveness of their teaching and learning practices. However, this value statement does not override the recognition that school-based data are multi-dimensional and contribute in a variety of ways to a more comprehensive, appreciative understanding of schools.

In the public forum, the focus on the use of systematically-collected empirical data in schools intentionally overshadows less formalized data sources like “anecdotal evidence.” Informal exchanges between teachers, undocumented observations of students’ classroom behavior, and affective descriptions of student achievement are all examples of anecdotal information regarded by many as a sub-class of data, i.e., second rate products of human perception. As detailed by teachers from Belleworth and Woodson below, however, these are all sources of data substantiating what some refer to as the “art of teaching” or the discretionary execution of education practitioners’ professional judgment. It is upon this very professional judgment that excellence in teaching and learning relies.

## Why Art?

### The Classroom Play-By-Play

In understanding the “art of teaching,” the first question to address is: What are the types of judgments education professionals need to make, or are expected to make as an element of their practice? Chapter 6 details the constant re-tooling of lessons Ms. Gavin engages in from period-to-period. She uses her knowledge of students’ individual strengths and weaknesses, as well as the pace and character of each class as a whole, to determine which instructional strategies she will need to best convey her content to different sets of learners.

Ms. Lovell described another typical moment requiring her to make on-the-spot instructional decisions:

*Okay so today we were doing... I was teaching with my math teacher, [and] we were doing this word problem that had about four parts to it. The teacher said, “Okay, we’re going to spend 10 minutes starting this problem.”*

*So a lot of the kids were getting the first part, and then... some kids read the second part, and then the teacher wanted to debrief.... So in that MOMENT there are four things you can do: you can debrief part A, you can debrief part B, but you can't debrief part A AND B because you're talking for 10 minutes and that's just too long, right? Or you can set up part C. You can just assume that everyone has done part B, and then go, “Okay, now part C.”*

*So in that moment, you can make these decisions.... So I have walked around, and I've noticed that a lot of kids have done A and a lot of them are starting B, but they're, like, confused... Like, half of the class has gotten this little part of B, and some of the class are just starting to read B. And then a handful of students have just finished A. So in that moment, I use my observational data knowing where the kids were at to know that I had to use this student's work –to project it on the board to debrief B and to propel them to C.*

*I think that's like a moment of decision, because a lot of teachers... Well I think any experienced teacher would just explain starting from A. This is the answer to A, this is the answer to B, this is the answer to C, how should we do D? But if you*

*just do D, the kids who haven't done B and C... they're not even listening to D. Then you just give the answer.... So those are the moves that teachers have to make on a day-to-day.*

In this excerpt, Ms. Lovell very specifically details a whole class approach to a math problem. In the consideration of how to collectively move learners through the activity, she finds herself needing to make an immediate decision about how to effectively scaffold the problem for students exhibiting slightly different learning paces. Ms. Lovell emphasized her use of “observational data” to assess where all of the students in the classroom are with their work. Noticing that not all students had progressed through the activity at the same rate, she opted to project a particular student’s work on the board as an example for the class, debriefing “Part B” and moving them forward to “Part C,” rather than simply walking through the problem from beginning to end. In this way, she could ensure that the entire class was able to review the material up to the point where most students had approached the work and allow the class additional time to continue through the problem (rather than simply giving them the answer).

In this decision, Ms. Lovell considered three additional approaches she could have taken, and her need to consider the length of teacher “talk time.” In the moment of the lesson, she and her co-teacher did not have much opportunity to confer or to ruminate on the way they ought to proceed. Instead, Ms. Lovell had to rely on her teaching experience, whatever information she had on her class at that very moment, and her professional judgment to calculate next steps. While such decisions may seem minute, these constant determinations of which instructional moves to take work in concert to build the momentum, pace, and fluidity of classroom learning. In this example, observational data on students’ class-based work, while not formally collected, recorded, analyzed, and interpreted, are an essential component of Ms. Lovell’s instructional practice.

## Impressions as Imprints

The need to make moment-by-moment decisions is one component of the professional judgment with which teachers are expected to approach their practice. In other areas, the overall impression that a teacher has of his or her class is an important element of diagnosing student strengths and needs in addressing the content. Mr. Urbina, from Woodson College Prep, discussed how he uses this data to develop an initial read on incoming students.

*You can glean a certain amount from just like classroom discussion and class participation. But that's more a... what's the word when it's data that's... based on like... um... it's just... anecdotal data. That's almost like anecdotal data because it's the impression that you kind of have.*

When asked how much that impression factored into his determination of how well his students were doing, Mr. Urbina replied:

*An English teacher once told me when I was a younger teacher, it was the beginning of the year, and I was like, how is it going? And he's like... "I hate the beginning of the year because you have to REALLY read everything." And I've always like... internalized that to mean, at the beginning of the year, you have to get to know your kids' reading, and writing, and thinking on paper. Because that is really... for the rest of the year... you can sense whether or not they're dipping or increasing, or if they're being... lazy on an assignment....*

*So I think that impression doesn't come through a rubric. It comes through almost like this fingerprint sense that you get for each kid. And so for me the verbal participation usually is a confirmation of things that I'm seeing on paper. OR... you'll see students who are very SILENT in class, but on paper are just like writing TONS and TONS. And... so like I said, it's not the-be-all-end-all. It's, like, ten percent. If I gave you a picture of a pie chart, it's like a ten percent sliver.*

Mr. Urbina makes an important distinction between obtaining a sense of student ability and progress through a formal rubric and developing a "fingerprint sense" of a student through classroom discussion and participation – what he deems "anecdotal data." Mr. Urbina argues that

it is difficult to develop an “impression” of a student in something like a rubric. Rather, in order to determine whether and how a student is exhibiting growth throughout the year, and to establish a sense of what engagement and effort from a student looks like on any given assignment, he relies on data from students’ “reading, and writing, and thinking on paper” in combination with their verbal participation in class.

These data are “anecdotal,” perhaps, because they are not documented for systematic review or calibrated to an objectively verified scale. But while these data may be considered lackluster in their empirical prowess, they seem also to be fundamental to excellent teaching. A teacher who is able to deduce his students’ individual ability, engagement, and potential through close reads of activity and work is likely preferred over one who relies solely on assessment scores to determine progress. This in-depth view to student achievement describes more than the use of “gut instinct,” assumption, or blind intuition, but the use of intimate, integral knowledge of a student’s approach to learning to inform instruction.

Still, Mr. Urbina readily admits that these kinds of information are imperfect, suggesting, for example, that a student’s verbal participation may not completely correspond with his/her level of writing, and vice versa. This suggests that “anecdotal” data are not wholly reliable in their depiction of student achievement (and why Mr. Urbina designates only 10% of his personal classroom data “pie chart” to these kinds of student observations). But while these sources of information are not the sole constitution of Mr. Urbina’s evaluation of student work and progress, they are still a crucial component. To extract a sense of a student’s reading, writing, and thinking style through observation, to develop an impression of an individual as a learner, is part of what distinguishes Mr. Urbina as a professional English teacher.

## Assessing Assessments

Teachers are also seen to exercise their professional judgment in evaluating the validity of external assessments and measures of student learning. Mr. Urbina went on to explain his department's use of a standardized assessment to determine a student's reading general level – another piece of data that he factors into his instruction:

*So... what else is in the pie chart is our reading assessment. That's another one where we're still kind of dubious of... how we feel about the assessment. Because... it gives us this kind of general lexile... estimation as to where [students] are at. BUT, I WOULD say that it does seem to jive with... the level of writing that you see in their work.*

Within this discrete example, Mr. Urbina touches on the issue of teacher validation of data sources used to make decisions regarding teaching and learning. Throughout this study, several origins of the mistrust of data were addressed, including teachers' limited technical knowledge of how data are derived and validated (see Chapter 6), lack of clarity as to how data will be used and fear of data misuse (see Chapter 9), and the encroachment of data-driven strategies on teacher autonomy (see Chapter 7). The point that Mr. Urbina raises here, however, is one which speaks to the importance of teacher professional opinion in performing a practical corroboration of assessment results with classroom experience. The expression of doubt over the reading assessment findings highlights an important reality check: Do the data appropriately correspond with what I observe in the classroom? While there is an arguable need for data users to be open to findings that do not reinforce previously-held beliefs, there is also a need for teachers to be critical consumers of data. Ensuring a reasonable level of correlation between assessment results and what teachers glean from student work is the use of professional judgment to determine the place technical instruments should have in classroom practice.

## Outside Opinion

While teacher professional judgment is used to validate external measures of student achievement, it is even more frequently used in informal assessments of student aptitude. Teacher-to-teacher exchanges of student behavior and progress were mentioned by several study participants as an important source of longitudinal student data. Mr. Nuñez, a teacher at Belleworth, highlighted the value of teacher comments included in a student's cumulative file in Chapter 6 in order to get a sense of past patterns of student behavior which might inform his current state of progress. Ms. Nava, also from Belleworth, saw her colleagues' professional opinions as important complements to her own appraisal of students. She explained how these exchanges are incorporated into her own practice:

*I'll give [my students] basic math problems, basic reading comprehension problems, and then if a kid doesn't do as well, then I'm like, oh hey, so like, can you tell me more about this kid? Like, who's had them? What were your experiences with them? Would you suggest that I work with them? And usually the teachers from the previous grade will say, "Oh you know," like, "you should keep an eye on so-and-so because they're doing really well," or "we need to make sure they keep doing really well." So usually it's like, I think it's more informally than... formally.*

Ms. Nava values her colleagues' professional opinions in providing essential contextual background about students she had herself distinguished as needing additional support. In this way, the opinions of her colleagues verify her own diagnostic work and a modicum of advice for whether and how those students might be further supported. Ms. Nava's passage also delineates that information is not always solicited, and her colleagues will actively spotlight those students who she should monitor and offer enrichment opportunities.

Mr. Urbina also discussed the role that his colleagues' opinions play in his evaluation of students' class performance:

*I don't teach the ninth and tenth graders here, I teach the juniors and seniors. So it's always funny because, after the first week, I will talk to the other teachers and go, "Oh my God, I love so and so! I love so and so!" And they're like... "Just wait." Or they'll be like, "OH REAAAALLY, that's interesting." They're like, "Has he written anything for you yet?" I'm like NO, we had a conversation about graffiti art. They're like, "Mhmmm (nodding). Wait 'til he has his first written assignment and tell me... tell me what you see." So that's kind of a meta-level of teachers conversing informally about kids.*

*And sharing stories of... Certain teachers... will give an assignment that provides insight that maybe your assignment DIDN'T, or your work HASN'T yet, or vice versa. And so... there are moments where we're kind of, conversing about students... you know... "Oh that person can't write. Oh that person really needs a lot of work around this." And then you'll come back, "Oh you know what? You're right." OR you'll say something like, "Actually we... you know, we tried this, we did THIS assignment, they really got engaged THIS way."*

In commending the value of his colleagues' professional opinions, Mr. Urbina pokes a bit of fun at himself for his occasional miscalculations of student strengths. First impressions of student personality, as an example, can sometimes be misleading as early student displays of thought and perspective are not always reflected in the writing they produce for class. Like Ms. Nava, Mr. Urbina looks to his colleagues' experiences as a valuable check against his own formative opinions of students. He also references a type of informal community of practice among teachers wherein instructional strategies are shared and discussed as a way of understanding student capacity. In this example, the needs of individual students, and the ways in which these needs are assessed and addressed, benefit from multiple perspectives. Assignments issued by another teacher may elicit different information about a student than Mr. Urbina's own assignments; Mr. Urbina may have been able to better engage a particular student through an alternative approach to the content.

Ms. Nava noted that she may experience differences in opinion with her colleagues, although they are not necessarily harmonized into her practice in the ways that Mr. Urbina mentioned:

*Ms. Gavin and I, we're really aligned with a lot of things, but I think, I don't want to say ed philosophy... I've noticed that the students that she's like, "This is an AMAZING kid," I'm just like, I can't stand this kid (laughs).*

*So um, and it's more of, I think it's the content area. So some kids are very much like English-oriented, like kind of artsy... very much like that, and when they come to my class, where it's more math-based and you know, more like practicing... And yes, we're doing labs, but it's more of, we're trying to get this answer, we're trying to develop new concepts, then they start kind of struggling.*

*And then students that I'm like, this kid is amazing, and then when they go into her class, it's kind of opposite. But I think it's just because the content areas are different. Other teachers it's really, we're kind of on the ball.*

Here, Ms. Nava notes a fairly consistent discordance with one of her colleagues, Ms. Gavin, with whom her own opinions of students do not align. She attributes this difference in perspective to differences in curricular content or, perhaps, their varying pedagogical approaches to the content. But for one reason or another, Ms. Nava finds that the type of students that excel in her own classes are not regarded as highly by Ms. Gavin, and vice versa. Ms. Gavin, while valued as a colleague, is not a reliable data source for Ms. Nava when it comes to characterizing students' classroom performance. Rather, she turns to other colleagues who maintain opinions more similar to her own for referral and conference.

In these examples, teachers are seen to use their professional judgment in making rapid decisions regarding their own instructional moves, gauging learner engagement and progress through curricular content, and in verifying appropriate measures of student achievement. Teachers regularly rely on the professional judgment of their colleagues in diagnosing and

assessing individual students' performance. Interestingly, all of these activities involve forms of data considered "anecdotal," i.e., undocumented student and classroom observation, subjective assessments of student work and classroom participation, personal valuations of student achievement measures, and informal exchanges of opinion with colleagues. The varying ways in which these data are gathered, documented, and exchanged are considered threats to their reliability. Data may be non-representative of student performance, for example, or inaccurately measured or overly prioritized by a teacher. All of these data sources are subject to falsification but rarely subjected to independent, objective experimentation.

And yet, at the same time, without the consideration of anecdotal data, teachers would be left with both a limited scope of understanding about teaching and learning and an unnecessarily restricted range of instructional moves. The ability of a teacher to flexibly adapt his or her lesson to the immediate needs of his or her classroom is an essential characteristic of responsive, receptive instruction. A teacher who is able to develop "impressions" of individual students' work is someone who attempts to understand their thought processes above and beyond what is conveyed by scores. A teacher is expected to be a critical consumer of data, as well as someone who actively engages in communities of practice and exchange with other teachers. In these examples, anecdotal data may not be the cornerstone upon which all judgments are formulated, but they do represent integral pieces of information in the exercise of perceptive, nuanced, and exacting teacher professional judgment. They give rise to teacher hypotheses about what is going on in their classrooms and different ways they might consider individual student need.

## Why Science?

There is a case to be made for a more methodical approach to teachers' collection of classroom-based data: a systemization of common teacher practice. In Chapter 7, Ms. Lovell considered her own instructional strategies and reflected on her desire to more explicitly document observations she makes in the classroom. She felt compelled to be more purposeful in her observation of student activity, challenging herself to answer the question, "What exactly am I looking for?" For Ms. Lovell, defining what is to be observed, documenting her observations, and then making sense of her findings in light of her overarching question is more than a personal interest in fine-tuning her own research acumen. Rather, this process is something she would more generally recommend:

*I would recommend it... to anyone. Because... I think teachers, it's easy for a teacher to just FEEL successful... based on student compliance. So if a group of students are AP level, whatever, and they're, like, compliant. It's easy to think that they understand, they're making progress, they're struggling with material, they're growing. It's easy.*

*So then you have to force yourself to look at certain things. Like what kind of responses are they actually writing? Just because they're writing a page of stuff, what are they ACTUALLY writing? Just because they're talking to their peers, like what are they ACTUALLY saying to each other? Like, what kind of vocabulary? I guess there's a lot you can look at.*

In this way, Ms. Lovell suggests that teacher practice can be improved by a directed investigation into student performance. She notes that it is "easy" for teachers to "feel successful" based on a general sense of classroom compliance. Ms. Lovell's answer indicates that, even when assignments seem to be going well and students appear to be producing work, it is important to scrutinize the quality and content of their work as a reflection of the teachers' learning objectives. In the case of classroom participation detailed in Chapter 7, Ms. Lovell

wonders if she thinks a lesson is going well because she sees hands raised, or if the lesson could be improved based on a more solid understanding of whose hands are raised and how often. More targeted investigation into what she observes during class would serve as an important examination of her own assumptions – inferences and beliefs she has developed over several years of classroom practice.

Also in an attempt to more systematically track student progress, Ms. Nava spent considerable time re-thinking her grading practices. She described how she is attempting to transition her grading system into one based on learning outcomes:

*I went to this really great training over the summer, where the guy just asked this question of, like, “Well, do your students know what they need to DO to pass your class, or what they need to LEARN to pass your class?” And that was just such a mind-shift for me because I don’t think I had ever been taught, or explained, how grading works.*

*It was like, you have to get grading done, but it was never... it was never this big conversation about like, what’s important to grade, what do grades really reflect? And so, I took that into the classroom and... I created a four-point scale... and I simplified it as much as I could for my students.... They know that if they have a 4 it’s [because] they not only understand it, but they can explain it to someone else and that person can get it. A 3 is, they get it, but they can’t really explain it. A 2 is... they kind of understand, but they need a little bit more support – they get stuck. And a 1 is like, I’m talking gibberish to them.*

*And so, I’ve been doing that a lot with them. And for almost anything I do that’s the grading scale that I have. And so it’s... those are now the pieces of data that I use. So everything from like presentations to writing assignments to even showing their work.*

Ms. Nava mentions here that her professional training did not include much detail on how grading “works” and, as a result, has been issuing grades to her students based on her own interpretation of how grading is conventionally executed. A summer training, however, made her reflect on these practices, causing her to question whether her grades reflected the work she

expected her students to complete or the content she expected them to learn. Ms. Nava then took a careful look at her curriculum, linking a new four-point grading scale to specific learning objectives she intended her students to master throughout the year, and stripping away the practice of assigning points for “creativity” or “timeliness.” In a later interview, she described how her students had subsequently developed fluency in the mathematical and scientific concepts they were meant to understand, as well as described their level of understanding in each domain using the 1-4 rating scale. In this way, Ms. Nava felt her grading was beginning to more accurately reflect her students’ understanding of the content. It enabled her to more effectively navigate through places in her curriculum wherein students need more assistance, as well as to empower her students in evaluating their own command of concepts and self-identifying content areas where they need additional support.

Both of these examples highlight ways in which teachers may more methodically approach the collection and analysis of classroom data in very practical ways. They do well to suggest that, in some cases, an “intuitive” approach to instruction may actually overlook the intricacies of what is being learned in a classroom, by whom, and in what ways. Systematically pursuing these questions is a way of “checking under the hood” of instructional practice rather than basing system performance on experience-based indicators of what seems to be “running smoothly.”

Still, it is important to maintain a balance in pushing schools to be introspective through more concerted data use and allowing them the flexibility to exercise professional judgment. Ms. Macia, Director of Autonomy and Accountability for LAUSD’s Pilot Schools, explained her stance on this tension from the perspective of District-led determinations of school performance:

*I believe that the general intent is to very much support our schools. I also think there is an intent to keeping the questions [of school performance] ambiguous, because in doing so, it does indeed allow for more discretion on the part of the directors and ultimately, the superintendent. Once you start writing things down and making criteria very specific, there's some good that can come out of that, but there's also some danger in that, in that... you become limited to what you have identified as measurable.*

*So, our district does a good job of getting ALL kinds of data. What we do with it I'm not quite sure sometimes but (laughs) we can go in and tell you how schools are doing with respect to the [state high school exit exam], with respect to graduation rates, with respect to a lot of different things, right? But there are also many things that we have yet to capture. How do you capture personalization? These are things that are foundational to pilot schools, right? [...]. How do you genuinely capture, accurately capture, parent engagement? These are things that we don't really measure.*

*And I believe that's sometimes why the question is left broad. Because there may be schools that are not performing, or meeting their benchmarks on these more, what to call them, more high stakes or popular, if you will, measures, but they may be doing fairly well in other areas. You would think that they are correlated, right? Who knows. You would THINK. But I don't know. I mean, I know what I don't know. I know that we don't capture this kind of data. How do you genuinely capture student interest? Some people would argue, well, if your attendance is high then they're probably engaged. Maybe. Not sure though.*

When asked about whether, in her experience as an instructional director, she felt that she had a general sense of whether a school had a good grasp on personalization, or engaging students, or being democratic and including voice, Ms. Macia answered:

*Yeah, I do think so. But again, it's very subjective. You know, it's... based on what Ms. Macia thinks based on Ms. Macia's experience, you know? And a lot of it, I don't want to say intangible things, but... how much joy is there on a campus, right? You know, does the principal LIKE doing what she does? Do the teachers like doing what they do, right? Do the kids' HANDS go up? Are they TALKING in class? Are they ENGAGED? And ENGAGED meaning... do they give eye contact to one another? Do they ask questions of one another? Do they ask questions of the teacher? What kinds of questions is the teacher asking? What kinds of work product do we see?*

*Those are the kinds of things that are difficult to capture in a way where we can make broader statements, accurate statements about what's really going on in a school. I think when you have people like me who have done it for a long time, you have this... a lot of anecdotal evidence that you can wrap in your head, but maybe not have it on paper. But then again, if you have five different directors, we'd all be looking for something different.*

Here, Ms. Macia makes a distinction between what is definitively measurable in assessing a school's performance and what is not. In the former instance, indicators of performance such as exit exam pass rates and graduation rates are examples of well-established and accepted metrics of minimum student competency. In the latter instance, a host of data sources are considered in the measurement of more abstract domains of school performance, such as "personalization," "student engagement," and "joy on campus." What Ms. Macia classifies as anecdotal data are those pieces of data that work together to form a general idea of a school, its environment, or the ways in which its teachers and students collaboratively engage in teaching and learning. But they also lack a level of accuracy and generalizability in their measure of "what's really going on in a school" and, as a result, are not officially recorded. These types of disparate data are informally collected, analyzed, and interpreted by individual District directors such that any resulting information is subjective and infused with personal perspective.

At the same time, Ms. Macia acknowledges the importance of anecdotal data in determining how well a school is doing. She emphasizes the necessity of the District to intentionally build some ambiguity around performance measures to allow for professional discretion. Discretionary judgments would make use of all data – including anecdotal data – that are critical in determining the health and strength of a school without hemming itself into only those characteristics considered "measurable." In this way, instructional directors working

directly with schools would be able to weigh “popular measures” of school performance against those considered valuable by educators, but perhaps less codified and systematically reviewed.

### **Cross-Case Insights**

These perspectives bring to light yet another variety of data educators use to inform their practice. While anecdotal data is recognized by many study participants as imperfect, it is also regarded as an essential component of understanding, in a well-rounded way, the multi-faceted and dynamic aspects of student, teacher, and school performance. Here we have seen instances where, in navigating complex notions of “effective” teaching and learning, educators are required to make sense of data where “measurable” outcomes have not yet been formally established. Charting a course through ill-defined situations that do not conform to a dictated set of rules, educators are regularly relied on to use their professional judgment in their evaluation of program effectiveness, student success, and actionable next steps, to name a few examples. Indeed, it is this very professional discretion that is sometimes intentionally protected by a level of ambiguity in defining expected outcomes.

This is not to say that anecdotal data should not be subject to some healthy skepticism. As Mr. Urbina noted, these data make up only a relatively small “piece of the pie.” As described by Ms. Lovell and Ms. Nava, there remains room for educators to improve upon their practices through a closer, more methodical approach to data use in the classroom. Ms. Lovell later commented that, while observation and intuition are necessary data sources for teachers, they are not sufficient and should be complemented by multiple data points to substantiate evidenced evaluations of student performance. There remains an onus of responsibility within the field of education to better define conceptual domains of practice and school effectiveness for purposes

of transparency, mutual understanding, and informed dialogue. Nevertheless, the significance of those data pieces that are less calculable should be recognized. This chapter contributes to a much-needed discussion around anecdotal data showing not only how and why educators need to exercise professional judgment, but that such professional judgment should be protected in the empirical assessment of student and school performance.

## **CHAPTER 9**

# **DATA FOR ORGANIZATIONAL LEARNING VS. DATA FOR ACCOUNTABILITY**

*I feel like schools are, like, little, you know, volcanoes of data. Because, there's just so much information and it's like flowing out at you all of the time. Right? And it's not just like numbers, you know, it's things like... even some of the uncountable, like you know, student comments, or parent comments, or things like that. So I think that there's just this, mountain, this huge mountain of data that exists, and so... I think in order to really use it is... It really has to start with like the people who are there, and the things that matter to them. You know what I mean?*

— Ms. Finche, Teacher, Woodson College Prep

Throughout this study, teachers have raised a host of school-based data types and sources used to inform teaching and learning practices in different ways. Among these are systematically- and unsystematically-collected data pieces; data required by the District, by the school, by teachers, parents, and students; quantitative measures, and qualitative feedback and responses; data aggregated to capture the achievement of large student populations and data specific to individual students' strengths and weaknesses; and, data on school-based interventions and whole school performance. As Ms. Finche states above, the endless “flow” of data streaming out of schools is analogous to a “volcano” of data-based activity.

Given the overwhelming presence of data, as pointed out in Chapter 4 what defines a school-based system of data for decision-making is less the variety and quantity of data available and more the intended use of that data. To reiterate Ms. Finche's salient advice, data use “really has to start with the people who are there, and the things that matter to them.” From this perspective, the challenge to schools is the wide range of stakeholders and stakeholder interests

that dictate data use – the varying perspectives, goals, and underlying motives that guide question development, the identification of data deemed to be relevant in response, and the ways in which data should be analyzed and interpreted in light of those overarching inquiries.

The guiding notion that data should be used for school-based decision-making encompasses a broad panel of purposes, some more explicitly stated than others. Borrowing from Patton’s categorization of the many intended uses of evaluation (2008), we might think broadly about the application of school data to the following five evaluative purposes: 1) *summative evaluation* (judging the overall effectiveness, merit, or worth of a program), 2) *formative evaluation* (using results to improve a program), 3) *accountability* (justifying or explaining what was done in response to oversight or compliance measures), 4) *knowledge generation* (looking across findings from different programs to identify general patterns of effectiveness), and 5) *developmental evaluation* (using results to change an intervention, adapt it to changed circumstances, and alter tactics based on emergent conditions).

Within this framework, the use of data to evaluate school quality, performance, and progress can take a variety of forms, each featuring its own distinct purpose. In education, it is not uncommon for these multiple purposes to be served by single sources of data. For example, the collection of student attendance is regarded in several different ways by different stakeholders. It is first and foremost a measure of *accountability*. Teachers and administrators readily discuss student attendance as a data piece that is “mandatory” or “required” to regularly report. Attendance is certainly seen as an accountability measure by the District which is liable for ensuring all eligible students are enrolled and actively attending school. But student attendance is also a key variable upon which pilot school funds are allocated – the higher the attendance rate, the more funding they receive. Attendance, as it is tied to the generation of

school revenue, can also be regarded as a *summative* indicator of a school's effectiveness. It may even be regarded as a comparative metric by which the overall effectiveness of pilot schools is evaluated. Still yet, for at least one principal within this study who noted that "the school's partnership with students isn't solidified until they're here," attendance is regarded as a *formative* measure of whether students were "buying into" the school culture. For individual teachers, attendance factors into the evaluation of a student's personal engagement in the content and their level of commitment to learning (not to mention their class grade). And for some, the utility of student attendance as a meaningful data source is marginal. As expressed by one principal, "You either have it or you don't." For many teachers struggling with the District's newly-launched education management information system, attendance records for individual students can be impossible to extract or are knowingly fraught with technical glitches. Given all these possible venues of data use for school performance evaluation, it is often unclear as to how school-based data are regarded and interpreted.

### **Perceptions of Data Misuse**

When the purpose of data use is obscured, unknown, or undetermined, so is the clarity of its political intentions. Several principals and teachers within this study often expressed frustration over the use of seemingly benign data for punitive purposes. Perhaps on one end of the extreme is the experience of Ms. Heredia, principal of Belleworth School of Arts and Technology. Her previous work as principal had been at a "small school," a District-endorsed reform model designed to better reach under-served communities. She explained that, after several years of operation, it was announced that having reviewed multiple data points, the District would revert to the original, comprehensive school model and collapse all of the small

schools into one high school. Ms. Heredia personally took issue with the “data” used to support this decision, particularly as she had been to a community presentation just three months prior wherein school-based data were used to exhibit the success of this particular small school initiative. She went on to describe the meeting at which the decision for collapsing the schools was announced. As she recounts, the official in charge defended the decision on the grounds that the data evidenced underperforming schools. Ms. Heredia raised her hand to ask what “data” were being referenced:

*Because... I know that our API scores, our reclassification data, our transfer data, our graduation data had all gone up. So what data was he talking about that looked bad?*

She went on to explain that the official simply replied, “CST [California Standards Test] scores.” Going on, she retold:

*I said, “CST scores? So, when you say that there are multiple data points that look bad, you’re really just using one data point – CST scores. Is that right?”*

Reiterating that the official confirmed that CST scores had gone down, she finished her account:

*I said, “OK. Just to be clear, it’s just the CST scores then.” And I said it just like that. I left it out there like that.”*

Ms. Heredia described how the official tried to provide additional explanation, but that she felt her point had been made. While she regards herself a “big fan” of using data and using it to make decisions, but she felt in this particular circumstance that the data had been manipulated; While data had been used to promote the small school partnership as an improvement over conventional models, at the last minute it was being used to justify closing her small school:

*It was ridiculous... They presented the same information backwards, like, we didn’t make enough gains. And it’s like, wait a minute, you just had a community*

*meeting to say that everything was working. Now you're using the same data [a different way].*

Ms. Heredia later discovered that the district official overseeing the small schools initiative was, in fact, running for Superintendent. This helped her contextualize his motives for focusing on the state exam scores. His “push” for the closure of the small schools based on CST scores was, in her opinion, because he was “going to be running for public office.” Ms. Heredia understood this to mean that the schools he was responsible for managing needed to show “positive” figures as evidence of his managerial success.

Although Ms. Heredia considers herself an advocate for the use of data for decision-making in school contexts, she felt herself witness to a particularly insidious use of her school’s performance metrics. The very same data that had been used to uphold the small school model was, just three months later, used to discredit the model’s viability. Long after the decision to disband the small schools had been made, Ms. Heredia now understood how seemingly “objective” measures of school performance could be used so discordantly. While the small school community may have been able to regard its performance data as encouraging, if the data were to be used to evidence a district manager’s personal “track record” of effective school improvement, standardized test scores would be the primary source of high-stakes decision-making.

Teachers also seem all too familiar with fluctuations in the ways in which classroom-based data are regarded. Heavy public debate, for example, surrounds the use of student standardized test scores – originally designed to measure student aptitude in specific content areas – in the evaluation of teacher performance, such as in the development of Value-Added Models. As another example, several teachers at Belleworth talked about the way data collected

from teacher observations have varied in their uses. Mr. Nuñez described his own take on classroom visits, which he considered par for course:

*I'm always thankful for my credential program. One of... that's what they taught us. It's your classroom, people come in, ignore them as if there is not, as if no one is there. You go through the motions, you set your tone, someone walks in, if you want to greet them, that's fine, but that is it. You focus on your students. I mean... that's what I do.*

Mr. Neal shared a similar perspective:

*I mean, when people come in and observe me, I don't get nervous, I don't panic, I just go. I gotta' just do what I do. You know?*

Still other teachers remember Belleworth's earliest days when classroom observations were far less inconsequential. When asked if he felt like there were "eyes on her" as a pilot school teacher, Ms. Gavin replied:

*I DID. When the school first opened, definitely. Because there was... I mean, it wasn't just feeling like that, there really WAS.*

When asked to elaborate on the Superintendent's observation of her class, Ms. Gavin went on:

*Yeah, and if he didn't like you, you knew it. And... if he felt you were doing something wrong, you were told in the middle of your class, in front of your class what you were doing wrong and that you had to change it.*

*So thank God he loved me. Some of my colleagues, no. So it was dreaded when he was walking through the halls. So the first three years, even last year, it still felt like that. It still felt like... we weren't good enough – you're not doing what we want, what you SHOULD be doing. But we didn't feel like we had guidance to show us, or to help us support.... This is the first year that I DON'T feel that.*

Throughout their careers, Belleworth's teachers have experienced classroom visits and observations for formal purposes of instructional rounds and Pilot School Reviews as well as for informal purposes, such as visits conducted by the principal, District administration, and the public. Frequency of exposure to visitors, and even formal teacher training, has imbued some

teachers with a sense of normalcy toward outsider observation. However, Ms. Gavin's recollection is expressive about the consequences of a "casual observation" turned into a public platform of scrutiny and criticism. Treating drop-in classroom visits as an opportunity to evaluate teacher performance during their lessons, the Superintendent instilled a sense of fear and "dread" among some Belleworth staff who felt as if they were consistently being told they were not doing what "they were supposed to." The residual sentiment among many teachers was that they "weren't good enough" and, at the same time, felt at a loss to make effective corrections to instruction without practical guidance and support. As a result, the conduct of classroom observations at Belleworth seemed to be approached with great caution by some teachers.

It is not just the reporting or collection of data that can serve multiple intentions and purposes; even the establishment of systems and structures for data use in schools can be caught in competing interests. Foxvalley School of Arts and Music has been operating as a pilot school for seven years and is considered one of the District's oldest such schools. Its current principal, Ms. Davila, previously worked as a district teaching and learning coordinator for the pilot schools and has been working with her faculty to establish a stronger culture of data use at Foxvalley. Despite her former ties to the District office, she expressed frustration at the heightened performance expectations mounted against pilot schools in exchange for their exercise of autonomy. In the following interview excerpt, she discussed how the District's view of pilot school "accountability" framed Foxvalley's work in a way she felt was both unpractical and unsustainable:

*So it's interesting. There's been a lot of talking... like messages that we get from... the District people saying, "Well you know, so you're going to have show how this is BETTER." And I'm always like, why does this have to be BETTER? You know? Like, we have to do as much or MORE if we want to have our OWN,*

*like, teacher evaluation system or something. It has to be as rigorous or MORE rigorous than the District one.*

*Or, you know... it's always like, "Well, if you want to make your own decisions, then you have to SHOW how you can be ACCOUNTABLE." So now there's more emphasis on ACCOUNTABILITY, you know? But it's not also letting us decide what WOULD be our accountability measures.... So we have this different model, but we still have to fit into the District system. So... it's kind of stressful actually.*

Ms. Davila describes how Foxvalley's innovative approaches to activities like teacher evaluation are treated as a "proof of concept" by the District and carry with their introduction the responsibility of proving such alternatives as "better" than District procedures. Ms. Davila's exasperation with this stems from several sources including the burden such evidence places on her small school with limited resources. She referenced a common pilot school grievance which points to the need for pilots to accomplish much more work than a conventional high school, including demonstrating its own "better than average" performance with far fewer personnel and financial resources. Moreover, Ms. Davila emphasized that Foxvalley did not have the flexibility to determine what measures of accountability would best apply to their innovative pilot school activities. She went on to say, in a later excerpt, that Foxvalley is held to the same performance measures as other conventional schools, such as exit exam pass rates and graduation rates. Not only are these indicators considered inadequate in capturing Foxvalley's innovative approach to practice, Ms. Davila also observed that these metrics are used to directly compare pilot school performance against the performance of all conventional schools, despite the differences in student populations that they serve.<sup>8</sup> In this way, the paradox of being a "different model" that still needs to "fit into the District system" feels unreasonable and unduly burdensome.

---

<sup>8</sup> Ms. Macia, the Director of Accountability and Autonomy for LAUSD Pilot Schools, has been working to change this system, this year producing reports of District-level statistics comparing schools within their respective "zone of

In each of these examples – Ms. Heredia’s experience of school performance data being used to both advocate for and subsequently suspend a small school initiative, the character of classroom observations at Belleworth dependent upon the intentions of different observers, and the frustration expressed by Ms. Davila in needing to constantly prove her school’s innovations as “better” than the norm – reveal a tension between the simultaneous use of school-based data for decisions related to both organizational learning and school accountability.

From the perspective of organizational learning, data are expected to be systematically collected, analyzed, and interpreted by people working within a school to construct a shared social meaning and to apply that information toward the continual improvement of the school (Louis, 2006). From the accountability perspective, schools are meant to respond to external expectations of performance using both description – What was achieved? – and explanation – How and why was it achieved and at the levels attained (Patton, 2008)? Data simultaneously used for both purposes carry a variety of perceived consequences with them. Confusion as to the purpose of data use can lead to stakeholders’ reticent “faith” in the ability of data to meaningfully improve school performance. For example, a teacher’s mistrust of student assessment data may stem from his or her own personal notion that such data are only collected to evidence the failure of schools in poor neighborhoods (a theory submitted by one teacher at The Academy), thereby impeding his/her inclination to feed assessment data into a genuine reflection on the quality of his own practice.

---

choice” or geographic areas within LAUSD comprised of multiple high school options, rather than against all of LAUSD’s high schools in general.

## Understanding Data in Context

One difficulty contributing to this tension is that teachers and principals see the importance of understanding their data in context – context that is not always considered when used for purposes of accountability. Ms. Lovell, at Woodson College Prep, characterized the utility of data in defining school performance in applying the metaphor of a beach:

*I feel like some people REALLY value... they think like data is everything... like the District or something.... The data is gonna' have the answers to everything, or... is gonna' somehow really show us clearly what the problem is. But it's SO NOT true. It's just NOT true.*

*Because there's so many things that we don't know. So, it's like how can you... OK, if there is like a beach of all this sand, right? You're at a beach, all this sand. And then when you collect data... collecting data is like collecting one grain of sand. That's how I feel. Collecting data is one little mark on a huge... experience of school life. And so how can you say that's the one thing that reflects the quality of the beach?*

When it comes to depicting whole school performance, Ms. Lovell feels that any given piece of data will be inherently limited. Data may offer some insight into school quality, or may assist in diagnosing a problem inhibiting improved performance, but a piece of data represents only one, singular aspect of “school life.” To assume that that datum will be entirely adequate in determining school quality ignores what remains undetected and unknown about the “experience” of education.

Providing a more specific example of why school-based data in context matters, Ms. Lam, a lead teacher at Foxvalley School of Arts and Music, and her colleague, Ms. Owen, discussed how year-to-year fluctuations in students’ exit exam scores had been a challenge to explain to the District. As teachers at Foxvalley, however, they have no problem in pinpointing the source of such seemingly dramatic changes:

*Ms. Lam: Last year we had a very, very high [exit exam] passing rate. But as part of a small school, pilot or otherwise, we are highly affected by... I don't want to use "caliber"...*

*Ms. Owen: Population shifts.*

*Ms. Lam: Yeah. Every year. And every year it changes. This year's senior class, we had something like 20% special ed when they came in as ninth graders. The 11th grade class, that's going to graduate next year, I had 10% gifted in that class that year.*

*Ms. Owen: You had 2%. Not even 2%.*

*Ms. Lam: Yeah. Four special ed kids that year.*

*Ms. Owen: So you see how that will throw it off? I mean... and of those 14-15 kids who have IEP's, over half are special day class students, which means they struggle academically more than [other] kids. You really can't expect them to pass it the first time.*

*Ms. Lam: Yet. (Laughs) So, that's it.*

Understanding that this situation would skew "the numbers" for her school size of 420,

Ms. Owen and Ms. Lam elaborated:

*Ms. Owen: And they'll say, "Oh there's something terribly wrong at Foxvalley." And the next year, we're the most brilliant school walking. And you just have to learn how to shrug it off.*

*Ms. Lam: Like last year we had over a 80% passing rate, and this year we dropped down to 70%.*

*Ms. Owen: You can put it right to those kids, even though they are awesome kids.... And I think you're held to a higher standard just by the statistical anomalies that happen when you have a small sample. So, no matter what, we're going to feel it occasionally just because... they won't control the population flow here. So, as a Special Ed Coordinator, I've got, you know, a group of 15 kids in a classroom where we have 25 kids..... I've changed the whole dynamic of the classroom. All of those kids need support, so how do I break them up over different periods?*

*Ms. Lam: And I also think which learners. Like Rainton School of Performance Arts [a neighboring Pilot school] DISPORPORTIONATELY gets the English learners. Right? So that affects their scores greatly. They have a FULL ELD 1 and 2 class, like something close to 30 kids. Our ELD 1 and 2 class is something like 15.*

*And that's not because we're cherry picking but, the way that the enrollment works is that whoever has an opening gets the kid. And so if we're capped at enrollment, then we don't get the kid, right? And then there are waves. Sometimes there's this weird WAVE of kids that come right around October. And then there's another wave that comes around April... We don't know when the kids are going to come. But they come in waves sometimes (laughs).*

For a small school with a population of just over 400 students, Ms. Lam and Ms. Owen emphasize that the constitution of their student body is more substantially affected by shifts in their student composition. Even fluctuations in relatively small numbers of students can represent larger percentage changes in Foxvalley's student body. In some years, Foxvalley's special education, or gifted and talented populations, represented a substantially sized group relative to the larger student body. In response, suggest Ms. Lam and Ms. Owen, Foxvalley's exit exam pass rate would appear to rise or fall rather drastically (by even 10 percentage points) from year to year. Likewise, so would the commendation and concern of the District. Rather than understanding the effect student population variables might have on pass rates, Ms. Lam went on to explain that one year they received a call from one of the District administrators to applaud their improvement inquiring as to how this was realized. Similarly, as Ms. Owen described, drops in performance were subject to appall and worry that "something [was] terribly wrong at Foxvalley." Together, they went on to explain the very technical nature of why and how student sub-groups are placed in their grades, classes, and even between the campuses in their multi-site complex. However, these are the minute details that are lost in a District-wide assessment of "what's working." When it comes to the District's use of exit exam pass rates to evaluate school

performance every year, Ms. Owen suggests, “you just have to learn to shrug it off.” The use of exit exam pass rates as a school performance indicator that does not take into consideration the detailed and sometimes complicated context of school populations is, at the end of the day, difficult to take seriously.

### **Practical Concerns, Conceptual Limitations**

Still another unintended consequence resulting from the use of the same data for both organizational learning and accountability purposes is that teachers may feel their own professional “worth” is poorly valued. In his reflection of Belleworth’s Pilot School Review, Mr. Neal advocated for a more meaningful classroom observation:

*And you know, I just wish that when we DID reviews for classes, it was a YEAR-long thing. Like, there was a person who sat in your class all year long to review and see EXACTLY HOW YOUR CLASS is in order, how it's being [run].*

*Because, you know, the 10-minute snapshots, or however long a person comes into your class, they're not getting the WHOLE picture. They're just getting a frame. If I'm not impressing you in a FRAME, you know, I could be doing great work here, you know? And you never know with that frame.*

From Mr. Neal’s perspective, to obtain an authentic sense of what teaching looks like within his classroom, he would want an observer present for an entire academic year. This period of time would allow an observer to develop an appropriately intimate view to his practice and to understand the nuances of teaching and learning as they occur. A 10-minute observation, on the other hand, will never capture the full scope of his work as a teacher. What it presents instead is only a single “frame,” and Mr. Neal has no control over what is observed and what is not observed in such a limited period of time. From a practical perspective, conducting a year-long observation of every teacher is untenable. But the point Mr. Neal raises is valid to his discussion

of data limitations. The “great work” Mr. Neal may be bringing to his classroom may go completely undetected in a 10-minute observation, translating into his concern that the data do not accurately represent his own performance. As a result, Mr. Neal finds it difficult to use the school review findings based on the classroom observation to reflect on his own instructional practice.

Ms. Lovell, too, expressed difficulty interpreting the value of her work through District measures of accountability. She explained in an interview how, just days before, she had received an email from the District monitoring the percentage of IEP reports she had submitted on time:

*I'm a special education teacher, so I feel like data is always used to... Like, you know I have to write these IEPs, and I'm the special education lead teacher, so they funnel through all of these data. The data is all about compliance. Like they meet these deadlines, how many am I overdue? How many am I on time? And that's the data that [are] supposed to reflect how GOOD my program is. Which then really UPSETS me. Because they are really looking for 100%, or an increase, I don't know....*

*It's not even the stuff ON the IEP, it's like whether or not I FINISHED an IEP on time. So it's the COMPLETION of an IEP. I know it's important because it's a legal document, but that's the only data that they look at as a special education teacher. As a special education lead teacher.....*

*I think it's aggravating because I feel like, the most important thing that I'm doing as a special ed lead teacher is... This last year we've launched co-teaching classrooms. So we went from segregated, self-contained classrooms, and we're including those students into most of their academic classes. And to me that is THE MOST MEANINGFUL work that I am doing.*

*And if I want anyone to judge my program, or my job, I want it to be based on THAT and not based on whether or not like I completed an IEP... ON TIME. And I understand it's like a legal document, and I understand that you know, it's... You know, the District bleeds money because of not being compliant on these IEPs, so I understand why. But if the one time I get an email from this person is for that*

*reason, then I feel like... a little grumpy about it. But... I understand. (Laughs) I don't think it's RIGHT.*

In this discussion, Ms. Lovell very clearly remarks that the data collected by the District with respect to her own practice as a special education lead teacher seems extraordinarily narrow. The extent of her work – some of it, such as full inclusion, difficultly implemented – is summarily reported back to her in the form of a compliance measure monitoring whether her team’s IEPs are completed and turned in on a timely basis. She mentions how “upsetting” this metric is as the sole form of “judgment” of her program and the quality of work her team is producing. While Ms. Lovell understands why the District needs to monitor the completion of IEPs for reasons of liability, this hardly explains how the timeliness of IEPs is a reasonable reflection of how “good” her special education program is.

An ensuing conversation with Ms. Lovell explored the use of school-based data for different purposes and, as she began to consider the idea that “different data are used for different avenues,” she felt that she was having a “mental breakthrough” with respect to her IEP data:

*But I’m really thinking like the more I talk... It’s kind of like cognitive therapy, but... I think what I realize is, to be clear about how data is being used might help, you know?[...] And then if we’re just CLEAR about that, then it would be less confusing. And then maybe less hurtful at times. Just to put like that data in context. I think that might help.*

*I really think I broke through! Like I’ve had some mental breakthrough! (Laughing) like, THAT’s why I get mad when I see those emails! And now if I just contextualize it in my mind in a different way, I will no longer take it personally! [...] Like, I will NO LONGER have it reflect my program.*

Here, Ms. Lovell begins to resolve the distinction between measures of District accountability and the kind of data that would, separately, more closely align with the

accomplishments of her program. She begins to recognize how the lack of clarity between these separate purposes resulted in her own personal confusion as to the intention and meaning of the IEP reporting data. As soon as she was able to contextualize the purpose of the IEP reporting data as an accountability measure rather than an evaluation of her program's effectiveness, merit, and/or worth, she was less inclined to take the compliance report data "personally."

### **Tainted Love**

In order to appropriately contextualize the District's reports on her IEP submissions, Ms. Lovell discovered the value of clarifying the purpose for which the IEP data were intended. Mulling this over some more, she reflected on the various types of data collected at Woodson and the multi-faceted reasons for which they were collected. She began to think about her own relationship to the data, and how an honest exchange between the data and users of the data could be contaminated by their need to serve multiple purposes. She suggested that there is a danger in co-opting data originally intended to inform organizational learning for accountability purposes.

An extreme example of how competing purposes may distort user relationships with data was provided by Ms. Heredia. In talking with about her former principal position, she described how she was witness to data manipulation by school staff:

*District policy was if a student's not on credits, they stayed at whatever grade level those credits are attached to. So what they started – and they wouldn't ask us of this publicly – they would have people come into our schools and kind of say... "This [way of hiding kids in the data] is a possibility." And so then, what ended up happening was we'd be under pressure as principals. Like, if WE don't do that, now it looks like our schools aren't performing. So are we going to do it or not, 'cause we're going to look like we're targeting ourselves –we're not going along with the bunch. So for example, District policy was move a kid down*

*to whatever credits they are. They wouldn't. Like, those schools would just keep those kids as 10th graders.*

When asked how the school would hide that in the data, Ms. Heredia went on:

*Well, 'cause the District wouldn't check whether the credits and the grade level aligned.... They gave you a deadline by they make sure you demote all your kids. And so, people just wouldn't. You know, they wouldn't demote them.*

*And so, you know, it got a little dicey in terms of, like, what people who were out of the classroom were seeing happening. Like people would send screenshots of stuff in the system, like student information to each other [laughs]. Like, "Hey, did they ask you guys to do this?" [...] Someone from that school would say, "Hey, I pulled up this data. Check out some shadiness that's going on." And the school's reporting all this growth... They would always share, like, "This school's doing better", but then we KNEW, like OK, 'cause we saw the screenshot of WHY they're doing [better] ... you know what I mean?!*

Not demoting students based on their credit standing, “rigging” suspension numbers, not offering certain courses so that a school receives a default minimum state test score rather than reporting scores lower than the default minimum, and pressuring teachers to pass students in the hopes of boosting graduation rates are all examples of school-based data “manipulation” cited by study participants. What Ms. Heredia describes is not only a propensity for some schools to show better figures in response to high-stakes accountability requests, but also to comply with surreptitious conventions of practice. Failure to do so would make it look as if her own school was not performing at the same standard as peer schools. But the production of data that was not, in all ways, an exact representation of her students’ performance would make it difficult for Ms. Heredia to determine what data were “valid,” and what data were “real in terms of what our teachers accomplished.” Ms. Heredia recognized that artificially inflating her school’s performance data would obscure her ability to determine the effect of teacher-led interventions and activities. Per her account, school staff flagged the potentially unethical nature of such data

practices and, while apprised of the data tricks employed by peer schools, did not report them to a higher level. As a result, Ms. Heredia maintained a covert understanding of the declared “growth” in some schools as an outcome achieved by misleading data rather than real, on-the-ground improvements in school practice.

While perhaps extreme, these examples show how the reliance on common data sources for both accountability and organizational learning purposes can lead to a manipulation of the data that renders them inaccurate for either purpose. The need for schools to evidence “performance,” “improvement,” or “growth,” presents a strong incentive for schools to create the appearance of gains. Performance data inflation, however, present imprecise details of school effectiveness to those concerned with accountability. Additionally, teachers and administrators are less able to infer real student growth related to intentional instructional changes from these data. The data subsequently lose utility informing the progress and direction of school practices.

Returning to Ms. Lovell’s reflection on her own, personal relationship with data serving multiple purposes, she described a much subtler interaction with data that led to outcomes similar to those described by Ms. Heredia. She provided the example of her use of reading data to inform her own practice while knowing these data were reported publicly as a measure of school performance:

*Like, once anyone publicizes a piece of data, and then you put it into the public, it becomes an accountability tool, and it taints the data. And it taints the way people USE it, and it taints the way people HANDLE it. So it's a very interesting thing.*

*So reading data, right? It's very important to my practice. It's very personal. I use it; it means A LOT to me. You know it's... it's what guides like 60% of my practice. So let's say a teacher does that, right?*

*But then that data... is collected.... Then somehow my... effectiveness as a teacher is being evaluated on that data, or the effectiveness of the school. Then it totally*

*taints it. Then it becomes... I think then it becomes weird. Like my relationship with that data number, it becomes weird.*

When it was posited that this occurred because the consequences were suddenly “high stakes,” Ms. Lovell replied:

*Yeah, it becomes high-stakes. And it becomes like... It's like I can't be as HONEST about this piece of data, or there's a... I don't know. And then there's this pressure to push the data. Then you make instructional moves that are not as helpful in the moment. Then you get obsessed with PACING, and then you get obsessed with like [laughs] the kids' deficits. It's like you're not MOVING fast enough. You're frustrated at the students.*

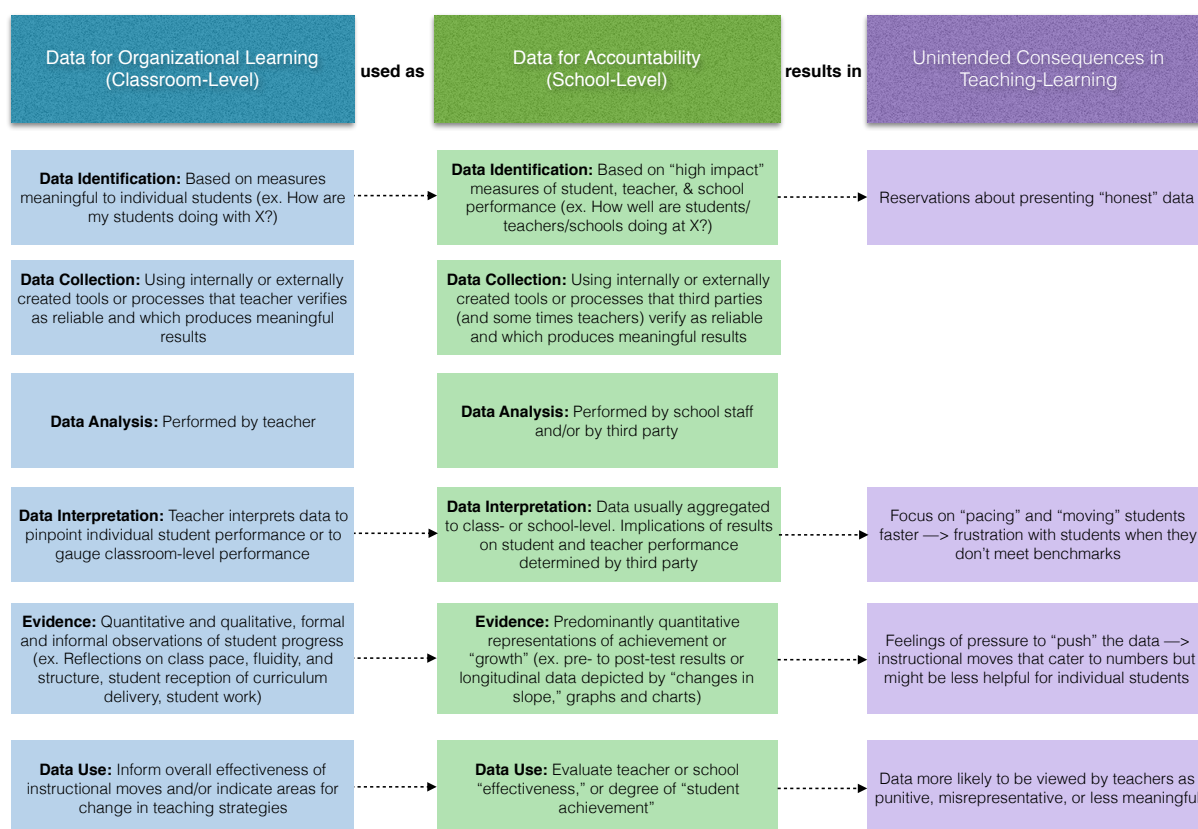
*And then if you tie that, not yet, but then if you tie that into teacher pay, then somehow it's like what you are incapable of doing. It's gonna', you know, not get you the bonus that I want...*

*Anyways, it's like that. It gets really interesting and really, like, messy. And I'm not saying that doesn't mean that it *SHOULDN'T* be publicized. You know I think we should be very... you know we're a public school, so it's public. But I guess, I guess just the reality is, that's where it gets messy.*

Here Ms. Lovell raises several important points concerning the treatment of data as it relates to classroom practice and to public measures of accountability. While she recognizes reading data as a primary resource in guiding the way she approaches her own students, once these data become publicized they take on a new meaning. At this juncture, reading data are no longer within the realm of control and consideration by the teacher, who might use this information developmentally, formatively, or diagnostically. Rather, they simultaneously pose as a higher-stakes measure of teacher effectiveness and student achievement. Ms. Lovell noted that her “honest” relationship with the data then became “tainted” as she feels pressured to then improve reported outcomes. A focus on moving the data (rather than individual students) leads to an emphasis on “pacing,” or the development of overarching instructional goals and strategies

that may not be as helpful to students in the minute-by-minute moments of learning. Frustration with students might be experienced when goals are not met. The interpretation of reading data takes on the feel of a summative evaluation of a teacher's performance. Particularly within high-stakes systems of teacher reward, this may inadvertently underscore a teacher's deficits and the feeling that he or she may be "incapable" of effecting desired progress.

In this way, data originally designed for teachers' instructional use loses its intimate connection with individual student progress and is gradually appropriated for external interests in accountability. The original intention of the data becomes co-opted into a separate stream of use for completely different evaluative purposes. This is not something that necessarily occurs at any one stage, but may occur across several stages of data use cycles (see Figure 3 below). As portrayed by Ms. Lovell, changes in regard toward the data may not be entirely conscious maneuvers, but could pose as natural reactions to positive and negative incentives.



*Figure 3: Challenges Associated With Multi-Purpose Data (A Teacher Perspective)*

Ms. Lovell's relationship with data, and data-based incentives, may not represent that of every teacher. Her depiction of fluctuating responses to different data purposes, however, showcases a relatable, almost economic perspective to data use in schools. Her delineated thought process speaks to the frustration expressed by principals and teachers throughout this chapter as they oversee and contribute to school-based data sets, catering to an assortment of data expectations.

This is not to say efforts to collect measures of accountability are nefarious in nature. As Ms. Lovell pointed out, such data is a public right. Additionally, it makes sense to capitalize on the use of meaningful measures of progress and improvement as an indicator of school and teacher effectiveness. Indeed, from a research and evaluation perspective, reducing the “burden” on school participants by using the same data sources for multiple purposes seems ideal. Schools are already inundated with enough work that the collection of additional data for varying purposes seems both impractical and unreasonable. Nevertheless, the unintentional consequences involved in re-purposing data must be recognized, as should be the roles of researchers, evaluators, and policymakers in clearly communicating their intentions, expectations, and use of school data from the outset.

### **Cross Case Insights**

Throughout this chapter, several examples of data misuse and misinterpretation have been provided by teacher and principal participants. Some of these detail intentional moves to use data in ways that support or defend particular political positions, such as in the case of Ms. Heredia’s former small school or Ms. Gavin’s experience of in-class evaluation by her previous superintendent. Still others discuss the false impressions data may relay when extracted from the context in which they were collected. Teachers from Foxvalley, for example, explained how apparent increases or decreases in school performance metrics were subject to fluctuations in their small school population. Ms. Lovell noted that a focus on data, which represent singular aspects of school functioning in brief moments of time, seems to distract from a more well-rounded understanding of the inherently complex quality of teaching and learning occurring within a school. Additionally, she noted that metrics observed at the school – or district-level –

for purposes of school performance accountability do not necessarily bear useful meaning in the context of classroom instruction. Expectations that these data are used to formatively improve instruction remain unfulfilled and feel out of place.

Finally, participants raised concern that how data are applied may persuade or dissuade the behavior of school practitioners in unintended ways. Ms. Lovell articulated how reading data, while substantially informative for her own approach to teaching, once published, introduces an inexplicit pressure for her to “move the data” as evidence of her students’ progress. In turn, her attention is shifted from individual student’s learning needs to whole-class strategies aimed at improving test performance. In an extension of their discussion over how changes in Foxvalley’s student body composition are known to have a calculable impact on high school exit exam pass rates, Ms. Lam and Ms. Owen added that, when the District calls to ask what work they have done to achieve seemingly excellent progress, the faculty feel inclined to detail Foxvalley’s programmatic improvements. The identification of population shifts as a primary source of data fluctuation is sidestepped in view of the need to demonstrate action and innovation.

The definition of “data” employed by this study considers data as information stripped of context – they do not have meaning in and of themselves. Data become information when they are connected to context and given meaning dependent upon an individual’s interpretation of the data. By this very definition, the meaning of data necessarily hinges on the purposes to which data are applied. While it is understood that data use is influenced by the perspectives, beliefs, and motivations of various school stakeholders, the experiences of teachers and administrators across Belleworth, Woodson, and Foxvalley provide essential insight into how competing uses for data can impact organizational self-concept as well as classroom-based teaching and learning activities. Even in the most well-intentioned circumstances, such as when formative school-

collected data are used as measures of performance to avoid additional the burden of data collection, these actions consequentially affect the way data are regarded by school practitioners. Wariness, distrust, or lack of buy-in to data use processes, then, can also be partially understood as a failure to convey, clarify, or clearly commit to transparent data use purposes, or simply result from the expectation that data can impartially serve all interests at once.

## **CHAPTER 10**

### **DISCUSSION**

#### **Introduction**

It is generally accepted that the use of school-based data to inform decisions around the development of student programming, improvements in instructional strategy, the allocation of resources, and the strengthening of student achievement is an essential characteristic of effective schools. The great variety of data collected from schools – governed by both the District and by schools themselves – is perceived to empower school stakeholders in their evaluation of performance. The ability to collect, analyze, and interpret data is also considered an opportunity to knowledgeably vary resource inputs, organizational systems and structures, and approaches to teaching and learning in ways that best fit school needs. Data are regarded as objective measures of effectiveness, allowing schools to empirically measure progress rather than yield to the erratic estimations of stakeholders inevitably colored by politics and perspective. Additionally, the use of data to track school improvements encourages stakeholders to articulate their goals and objectives, and to develop activities specifically designed to meet those targets. Data are viewed to promote reflective strategic development in schools and discourage the outright abandonment of interventions and innovations in exchange for documented cycles of iterative improvement. Schools that use data well are therefore regarded to manage themselves well, and schools that manage themselves well serve students better.

While these perspectives are readily accepted on a policy level, there was a need for a more in-depth and nuanced understanding about what data use processes look like in the day-to-day context of schools. Whether schools feel adequately supported in processing data for use in decision-making, how data are actually incorporated into deliberation and discussion, and how

we think about what characterizes “effective data use” at the school-level are all outstanding issues of interest in the widespread promotion of data use in schools for decision-making. In response, this study closely examined processes of data use in three pilot high schools in the Los Angeles Unified School District in a comparative case study. This approach emphasizes the unique frame of reference each school brings to data use in everyday implementation. It acknowledges that the meanings of “data,” and the ways in which data use processes are undertaken, are actively constructed by schools and naturally varying. In this way, data use is regarded as culturally-defined.

### **Understanding and Supporting Data Use as a Part of School Culture**

This study produces a picture of data use within schools that is complicated, context-dependent, and in constant fluctuation. The recognition of data use as a cultural process suggests that schools’ successful use of data to inform decision-making is not nearly straightforward. Thus, in answer to the research questions guiding this study asking what school practitioners identify as credible data, how data are used to inform decisions related to school improvement, strategic planning, and instruction, as well as how data are used to monitor school performance, it was found that the cultural and organizational characteristics of schools shape the ways in which teachers and principals use data for all of these purposes. The active use of data is not only contingent upon what are accessible by schools, but also by how users value and prioritize various data sources, as well as their experiential and technical knowledge in practically applying data to questions of instruction, strategic planning, and school-wide improvement. These skills are largely influenced by stakeholders’ perceived sense of agency in decision-making processes, the collection and synthesis of multiple data sources (including anecdotal, observational, and

systematically-collected performance data), and approaches to data collection, review, and analysis that offer teachers personal and professional support alongside continuous cycles of development, piloting, and application. In short, data use within schools is not nearly systematic. Rather, what this study has shown is that the ways data come into play in various decision-making moments varies substantially across school sites, as well as within each school case.

The idea that school-based data use is culturally dependent also yields to the notion that data use processes cannot simply be transplanted from one school to another with the same effect. As such, it is difficult to identify strict indicators of “schools that use data for decision-making effectively.” In response to the question “What does a school that uses data well look like?” there is no definitive list of components or characteristics. It would be insufficient to suggest, for example, that a school should be reviewing its data three times per semester, or that each school must establish a data team tasked with the responsibility of analyzing, interpreting, and disseminating data to faculty. Such guidelines would not respect the exclusive relationships schools necessarily maintain with data use processes.

Certainly there are approaches to data use that have been positive for many schools and which serve well as overarching guidelines for practice. This study supports previous research indicating that when schools have reserved time, financial resources, and human resources (not just in name, but as a considered investment), and wherein schools have the infrastructure to access and compile data, they are better positioned to leverage data to inform their decisions. This difference was seen, for example, in the comparison of Woodson College Prep, which had invested several years and substantial financial resources into the organization, synthesis, and collection of data for use by teachers and administrators, and The Academy, which relied heavily upon one or two members of their faculty to access, extract, and analyze District or student

survey data during whatever time they found available. This study has also shown that data appear to be more effectively used for purposes of decision-making when they are routinely assembled and organized in response to “research questions” closely aligned with stakeholder interests, wherein schools have systematized processes of decision-making that are respected by faculty and administration, wherein stakeholders feel their voices are honored in decision-making processes, and when methods of data collection and analysis make practical sense to the stakeholders responsible for using those data. Another comparison highlighted within this study was the perceived value of test facilitation, scoring, and analyses of Woodson’s “common assessments” expressed by teachers from the English and Science departments against the experience of one faculty member from the Social Studies department, who struggled in her endorsement of her own test as a credible representation of student capacity. Data and data use processes are more likely to be valued when school stakeholders understand the purposes underlying data use routines, when the intended uses for data are transparently communicated and observed, and when teachers are deeply involved and feel that processes of data identification, collection, analysis, and interpretation are in-step with instruction. The activity led by Belleworth’s principal and Instructional Leadership Team to physically represent the number of students failing each teachers’ class was an example of the importance of connecting teachers to student performance data and the resulting implications on classroom practice.

Data are found to be most useful when they are regarded not just as “numbers” or as the preeminent determination of school performance. Rather, data must be understood in context. It must be recognized that there exist limitations to data in their portrayal of student, teacher, and school achievement, as well as naturally occurring threats to stakeholders’ own processes of reflection and self-critique. What school stakeholders see in the data is influenced by their

knowledge of the conditions underlying data identification and collection, as well as the varying perspectives they bring to data analysis and interpretation. As such, data are not components of a bounded system of rationality, wherein decision-makers synthesize, prioritize, and determine action based on transparent estimations of costs and benefits. By definition, without context, data are uninterpretable – they are conferred meaning based on cultural and contextual factors such as purpose, place, and time. Data are not outcomes in and of themselves. While data do not necessarily convey rationalistic value they are, however, better thought of as tools to guide the way we speak about our schools. Albeit abstract, thinking about data and the ways we attempt to make use of data in this way helps us to characterize our approach to data use in school settings. It helps us to reflect on the question: Are our immediate and long-term expectations for data use in school-based decision-making reasonable?

### **Re-thinking Data Use for Decision-Making in Schools: A Revised Conceptual Framework**

Further deliberation around the question of how we think about data use in schools may be guided by the conceptual map presented in Figure 4. This framework is built from the perspective of schools as to what practices encourage effective data use. It attempts to outline what major factors schools and school leaders should acknowledge in determining how to best support the use of data in processes of decision-making. While based on the Coburn and Turner (2011) framework of data use in schools, this framework substantially revises the portrayal of data use as a schema of nested and intertwined relationships between individual stakeholders, organizational systems, and data use processes. There are no inputs and outputs – an intentional statement against data use as a linear procedure fed by resources and interventions and generating improved school management, instruction, and policy as outcomes.

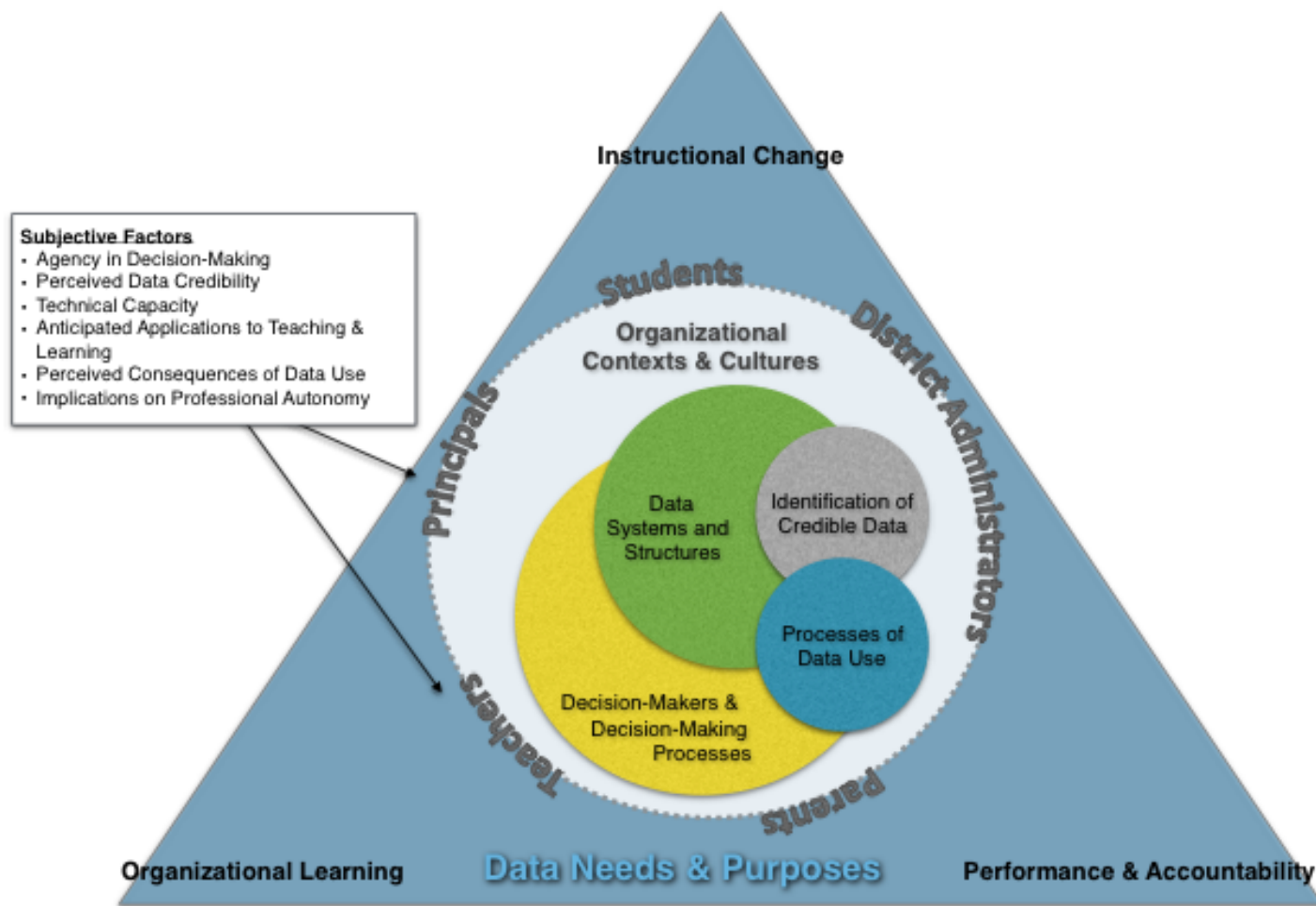


Figure 4: Framework for Data Use in School Decision-Making

## Data Needs and Purposes

Organizational learning, school accountability, and improved instruction are seen as the primary *purposes* driving data use in schools. These purposes are the background against which data use processes are developed, and set the stage for how data use processes are implemented. Constituting the two points at the base of the Data Needs and Purposes triangle are *Organizational Learning*, which asks the question *How well are we doing with respect to our self-determined standards of success?* and *Performance & Accountability*, which attends to the question *How well are we doing with respect to external expectations of performance?* While these two ambitions of data use in schools are distinct, they are often found to be in direct conflict when cast as objectives that can be simultaneously achieved through the same data use activities. The question underlying *Instructional Change* as a driving purpose of data use asks *How, if at all, do changes in classroom instruction affect student learning and in what ways?* Although rarely stated as such, improvements in Organizational Learning and school Performance & Accountability both imply changes in instruction. Organizational Learning is, therefore, placed at the top of the triangle in recognition of what is tacitly expected from all school-based data use activities and as an occasionally explicit purpose of targeted data use interventions.

The different purposes of data use are portrayed as underlying – if not directing – data use processes within schools. Interventions, tools, and policies aimed at encouraging data use each bear their own intentions. Each, therefore, brings with it a different purpose for data use. Furthermore, school-based data needs are in constant flux alongside the development and revision of accountability measures, organizational improvement, and instructional change. It is

not always anticipated what types of data will be needed, even when decisions feel fairly clear-cut and there are readily identifiable decision-makers (Weiss, 1988). However, when we change the purposes to which the data are applied, we change the nature of the data. Too often we fail to recognize how each purpose influences how data are construed, valued, and used. As a result, we can also overlook the unintended consequences of re-purposing data for multiple purposes, such as data misuse, non-use, or the construction of misleading data.

### **Stakeholder Perspectives**

Within the expectations of data use for various needs and purposes, the ways in which data are used in schools is determined by the organizational context and cultures of schools and comprised of several systems and processes. Guiding and governing the contexts and cultures of each school are school stakeholders, particularly District administrators, principals, teachers, parents, and students. These perspectives are integral to each of the contextual components surrounding data use and are shown as encircling the organizational contexts and culture of data use. Each stakeholder group experiences teaching and learning processes differently within a school and, as a result, each is expected to hold an independent view of its systems of data use. In considering the four domains comprising the “Organization Contexts & Cultures” of data use, it is essential to recognize similarities and differences between stakeholder orientations. Importantly, while stakeholders are presented as categorical groups within this framework, this study has emphasized that individual perspectives play a large factor in school-based data use. We know that the beliefs, values, and assumptions held by individuals substantially impact what is observed or not observed in the data, or what is eventually defined as credible (Donaldson,

Christie, & Mark, 2014; Young & Kim, 2010). School leaders should take care not to assume, for example, that all teachers hold similar opinions about data use, or have the same capacity, experience, or motivations in their customary use of data.

### **Decision-Makers and Decision-Making Processes**

Looking at how data use might be regarded as an artifact of school culture, emphasis is placed on the central role decision-makers and decision-making processes take in defining data use. This study has shown that how decisions are made within a school, and by whom, must be clearly understood and accepted by school stakeholders before data are able to enter into the conversation. This domain is thus depicted as underlying all others within schools' organizational contexts and cultures. As depicted in Figure 4, the general population of school stakeholders is illustrated as distinct from school decision-makers. Decision-makers are those individuals designated with the authority and responsibility to institute changes within the school. Decision-makers are likely to be school leaders, but may also have roles outside of conventionally defined positions of "school leadership," such as principals and other administrators. Who the decision-makers are in any given situation is also dependent upon what decisions are being made. The designation of decision-makers should not be confused with the endorsement of individual figures of authority. It should be remembered that the decision-making processes detailed within this study have focused on the role of teachers as an active decision-making body. As detailed by the work of Park and Datnow (2009), teachers see themselves as "knowledge brokers" and consider it a duty to connect with one another to exchange knowledge and expertise. The authors submit that, without collaboration and

collegiality, data is impossible (Park & Datnow, 2009). Alongside the determination of who will be charged with making final decisions, then, there is also the consideration of how to ensure the voices of multiple stakeholder groups are effectively heard and incorporated into decision-making processes as first steps towards imbuing a sense of collective value for data and data use.

### **Data Systems and Structures**

Formalized school systems and policies to routinely review, discuss, and disseminate data, as well as the infrastructure to compile and analyze data from multiple sources enable effective data use. They are not, however, found to be mandatory – teachers and administrators were discovered to have made use of disparate data sources when formal systems and structures were not in place. However, study participants had been better able to leverage data after taking stock of what data were available and accessible, assembling the data deemed most appropriate for review, and routinizing regular opportunities for data analysis and interpretation. The development of these systems and structures are motivated by both internal and external data needs. It inherently involves the determination of what data are considered a priority for a school – often a negotiation among school stakeholders. The ability to respond to data needs and requests is supported by the establishment of data use systems and structures. Corroborating prior research (Lachat & Smith, 2005), it has also been found within this study that the integration of data use systems and structures into school routines is best facilitated when they are built to respond to the expressed needs of decision-makers and in ways that are accurate, consistent, and timely. Thus, in Figure 4, data systems and structures are depicted as embedded within the decision-making domain and as an important foundation for other data use domains.

## **The Identification of Credible Data**

Part and parcel of the systems, policies, and processes discussed so far is the identification of which data a school should target its focus on. On the one hand, schools within this study have emphasized that there exists an overwhelming amount of data from which relevant data must be selected for use in school-based decision-making. On the other hand, teacher participants have frequently highlighted an absence of systematically-collected data which accurately capture the experience of teaching and learning. Somewhere in between, school stakeholders must determine which data they view as credible in decision-making and for which purposes. That is, schools must actively decide what data are considered practically useful in making school-based decisions, are relevant to practice, and are valid and reliable reflections of student, teacher, and school performance. This conversation is heavily influenced by the kinds of decisions to be made and who is making them, as well as what data are, or are not, readily available to schools.

The process of identifying credible data is seen, in some ways, as separate from more general processes of data use. Data may be handled by a school in response to external requests as a matter of compliance. But data that are effectively used by schools in processes of decision-making – data that even make it to the table for review and discussion – must first be acknowledged as “credible” (Donaldson et al., 2014). As seen within this study, this is heavily influenced by stakeholder perspective. Alongside the determination of data credibility on a whole-school basis, individuals also determine for themselves what data are credible. Interestingly, personally differentiating views on what constitutes credible data – even when they substantially contrast with a more collective sense of credibility – may not derail school-wide

processes of data use. Traditions of teacher autonomy can work to compartmentalize differing approaches to data use. Data may be used relatively effectively to inform long-term strategies or school-wide programs and interventions, for example, even where individual teachers may not use those data to inform their own classroom progress. Nevertheless, whether at the school- or individual-level, what data are identified as credible influences how data are used in decision-making moments, and this domain is seen to overlap with the “processes of data use” domain in Figure 4.

### **Organizational and Individual Processes of Data Use**

As suggested by Coburn and Turner (2011), processes of data use involve a number of activities related to the psychological processing of data, such as the noticing of patterns and trends, data interpretation, and the construction of what implications data have on school-based decisions. Like the identification of credible data, school-wide data use processes are seen to draw on the beliefs, motivations, and knowledge of individuals, as well as the social interactions individuals have with one another. Research suggests that stakeholders tend to notice in the data only what reinforces their previously-held beliefs, assumptions, and experiences, and filter out data that might contradict or challenge these beliefs (Bickel & Cooley, 1985; David, 1981; Hannaway, 1989; Ingram et al., 2004; Kennedy, 1982; Young & Kim, 2010). This study has found some examples of this. More predominantly, however, it was observed that teacher participants felt challenged by their limited technical capacity to analyze the data and then draw meaningful connections between presentations of data and the teaching and learning activities taking place in their classroom. Furthermore, teacher participants expressed difficulty moving

from the identification of student needs in the data, to developing actionable next steps to address those needs, and then converting the knowledge gained through data interpretation into student learning. Thus, alongside the need to challenge teachers in taking a constructively critical approach to their instructional practices by looking at data, there is also the need to acknowledge teacher concerns around the limitations of data, as well as understanding what data mean in their classroom and school contexts.

Recognizing individual teachers as primary agents of data use in schools adds to the Coburn and Turner (2011) perspective of data use routines which they define as “the recurrent and patterned interaction that guides how people engage with each other and data in the course of their work” (p. 181). This definition suggests that data use is necessarily a focused activity involving multiple people and does not yet discuss individuals’ routine use of data independently. This is particularly relevant within schools where teachers have been observed to draw on multiple sources of classroom-based data to independently inform immediate instructional moves, evaluate student performance, and revise and refine longer term pedagogical approaches. Principals have also been seen to conduct their own analyses of data as a way of guiding school strategy, constructing agendas, and identifying issues for whole-school input. In these contexts, individuals within schools task themselves with the responsibility of organizing, analyzing, and interpreting data they consider credible in response to their own professional needs. This adds another layer of complexity in discussing data use within schools because individual data use practices may or may not be in step with whole-school data use routines. Additionally, as individuals’ technical understanding of data and data use processes are diverse within schools, their experiences working with data may positively or negatively reinforce

independently-held motivations, beliefs, and knowledge about data – personal aspects known to influence collective determinations of data use. These tensions give credit to the notion that schools as organizations are not just the representation of collective interests. Rather, they suggest that the Coburn and Turner (2011) definition of data use routines - organized around specific interests and goals - insufficiently captures the role individual interests and actions have to play in manifesting a school's behavior around data.

Previous research on faculty-led data use processes has underscored the importance of cultivating trust and collaboration among teachers as the basis for constructive conversations around data and their implications on practice. It is argued that conventional norms of privacy inhibit the ability of teachers to talk in depth about their instruction and share evidence of student learning with their colleagues (Little, 2007; Little, Gearhart, Curry, & Kafka, 2003b). Indeed, a reticence among some teachers to share student outcome data or details of their instructional practices with colleagues was encountered in this study. However, barriers to transparent dialogue were not seen to simply result from an unwillingness or a sense of professional privacy. Rather, trust among colleagues – and especially between teachers and administrators – was seen to be the result of efforts to form strong personal relationships with one another. Additionally, the use of data as a tool to talk analytically about teaching and learning was observed to rely on multiple subjective factors including individuals' feelings that:

1. What I have to say about data is valued by decision-makers, and that I have a stake in decision-making processes.
2. The data are credible and trustworthy in their reflection of my classroom practices and what my students have learned as a result of their work with me.
3. The data are aligned with teaching and learning outcomes I value.

4. I understand what the data do and do not represent from a technical perspective.
5. Looking at data is an opportunity to show where my students are and identify ways I can more effectively reach them, rather than a chance to disparage me for what my students and I are not doing.
6. I understand how the data may be helpful in understanding student/class/school progress and may be useful in encouraging further improvement.
7. Even though we might collaboratively determine preferred approaches to teaching and learning as a faculty, it is understood that flexibility is sometimes warranted and that I maintain the right to determine when to exercise that flexibility in my professional space.

That is, influencing feelings of vulnerability and trust amongst colleagues were teachers' perceptions of agency in decision-making processes, their own facility in understanding and interpreting data, perceived consequences resulting from data discussions, the ability to directly apply data to teaching and learning activities, and the ability to maintain professional autonomy. Based on this study's findings, these elements are represented in the revised conceptual framework as individual-level variables in relation to teachers and principals. Further research is encouraged in the exploration of individual-level factors influencing data use for groups of school stakeholders other than teachers and principals.

### **Practical Applications of a New Theoretical Approach**

This revised conceptual framework presents a new perspective on how we think about data use in school contexts. While seemingly theoretical, this framework also presents a practical way of thinking about how effective data use might be best supported in classroom and school contexts. To facilitate this function, Appendix B provides a list of important considerations school leaders interested in bolstering data use might find helpful in reflecting upon, and

identifying areas of improvement in data use processes. Rather than being prescriptive suggestions, these issues are raised as “guiding questions,” the answers to which are necessarily characterized by each school’s unique context and culture.

### **Lessons Learned**

The ability to effectively apply school-based data to decisions around teaching and learning in schools is a careful balance between the needs, expectations, and values for data held by individual stakeholders and by the school as an organizational culture. These factors are necessarily driven by the specific contexts in which data use is meant to take place. While the needs of each school differ, data experiences among participants have drawn attention to some key areas of future focus in support of data use that are worth reiterating.

### **The Myth of Data Transparency**

Frustration with data use processes often stemmed from a lack of transparency around the intended purposes of data, as well as the technical aspects surrounding data collection, analysis, and interpretation. Ironically, although data are purported to be a transparency tool serving as objective statements of student and school performance, there is much confusion around what data do, and do not, represent, and the extent to which interpretations drawn from data are valid. Data, as it turns out, are not entirely straightforward. While this may be so, the expectation that, as one data-savvy participant put it, “The data speak for themselves, don’t they?” prevails. Pilot schools are expected to refer to data as a test of their innovative approaches to teaching and learning, teachers are expected to conduct classroom pedagogical experiments and to translate data into improved learning outcomes, and schools are expected to respond to both

accountability data requests as well as to the public reactions resulting from those data. However, within each data-based activity, there seems to be substantial variation among participants as to their knowledge of data collection and analysis methods, as well as how to make appropriate inferences from the data. These findings echo previous understandings of the variability of data interpretation dependent upon a person's or organization's existing beliefs, values, and norms (Cronbach et al., 1985). What is noticed in a school environment, whether that information is understood as evidence pertaining to some problem, and how it is eventually used in practice is regarded to be reliant upon the cognitions of teachers and administrators operating within a school (Spillane & Miele, 2007). This study continues to find that differences in the ways that school practitioners interpret and use data are also, in part, due to a need for increased technical capacity – the conversion of teachers into in-class educational researchers is reliant upon specific training regarding the technological manipulation of data, statistical analysis methods, and evaluation strategies (see “Professional Development” below).

The responsibility to convey data use processes in ways that are transparent and understandable by multiple school stakeholders (and especially teachers) also falls on facilitators of data use activities. Often as a result of limited time, the methods guiding data collection and analysis remain undisclosed. Teachers are frequently asked to look at assessment or evaluation results during stages of interpretation without a proper introduction to the methodological choices underlying those data. As a result, there is a substantial amount of confusion around why, for example, specific items were selected to comprise an assessment or evaluation, how often data were collected, when, by whom, and in what manner, what statistical procedures were used to compile and analyze the data and how these affect the interpretation of results.

Additionally, insufficient knowledge regarding how the data are going to be used undermines participants' confidence in the integrity of the data. This is particularly true when the same data sources are tasked to address multiple data needs – while some purposes for data collection may be accepted, others may bear unknown risks.

While it is certainly recognized that time is on short supply within schools, data users are entitled to some declaration of the research and evaluation methods guiding processes of data use. Open discussions regarding perceived benefits and risks of the data would contribute to data users' understanding of what motivations prompt data use processes, dispelling (often silent) concerns that data will be used insidiously or for unintended purposes. Walking through data collection and analysis procedures in this way can feel time consuming, but investments in school members' thorough understanding of the techniques, methods, and motivations underlying data use activities is also an investment in their propensity to buy-in to, and effectively use, those data.

### **Data Used in Decision-Making Are Part of the Process, Not the Outcome**

In the same way that school stakeholders are expected to better adhere to research and evaluation guidelines in the examination of their own practice with data, the research, evaluation, and policy community must be more forthright in acknowledging the limitations of characterizing school practice with data. As advocates of data use, we are sometimes so focused on promoting data that we neglect to talk about what data do not do. Herein lies a philosophical strain with school stakeholders who know that data are only one component of a story. While data may present unique insight into teaching and learning, and serve as an important tool for

dialogue and discussion around school improvement, they do not constitute the totality of our knowledge about schools or what happens in classrooms.

Spillane (2012) provides a helpful theoretical frame in understanding this tension, suggesting an “ostensive aspect” and a “performative aspect” from which we might research data use in schools. From the ostensive aspect, he argues, our attention is directed to formal organizational structures that school leaders, policy makers, and reformers use as a vehicle for changing everyday practice. From the performative aspect, we focus on practice as a central concern in investigations of data production and use. The performative aspect also gives credence to the notion that practice unfolds over time, that practice is the outcome of interactions among school stakeholders rather than just individual actions, and that situations serve as a medium for those interactions as a defining aspect of practice.

Ostensibly, then, it is understood by participant schools that data use practices are an essential component of good practice. But even with the institution of data use routines within schools, from a performative aspect, teachers and researchers continue to grapple with how to evidence causal relationships between instructional change and student learning *in situ*, or how to capture a more holistic picture of student growth that includes both behavioral and academic characteristics. In examining the use of data to inform instruction, members of the education community argue that there are times when the use of professional judgment is warranted, that the context of data matter, and that the many types of data education professionals draw on (even if anecdotal or unsystematically-collected) have value. This is not to say that systematized processes of data use have no place in schools; that what data are considered credible to different educational stakeholders is an on-going, and necessary discussion.

In practice, this implies that school-based data should not be regarded as outcomes in and of themselves but rather as indicators of performance. It is imperative that, in tracking performance benchmarks, data are understood as they relate to overarching goals and objectives. For example, is it important that exit exam pass rates have improved by 10 percentage points or that staff attendance rates reached 90% for the year? Or is it important that high school students are adequately prepared for entry into the workforce or an institution of higher education and that staff are committed to and engaged in their work with students? To what extent do data serve the estimation of these more conceptual domains of practice, and what additional data must be considered to inform the larger picture? This is not to say that numerical representations of practice have no place in the evaluation of student, teacher, and school performance, but perhaps that they are necessarily insufficient in understanding the great depth and complexity of teaching and learning. The review of school performance metrics presents an important opportunity to empirically examine progress. The great effort to internalize data and data use routines into everyday practice, as well as the significant technical demands data place on schools, however, can result in an emphasis on data patterns rather than what they are meant to portray.

### **How Data Are Not Used**

One substantial argument for the promotion of data use in schools is to better enable faculty and administration in monitoring the progress of specific student subgroups. Performance data are perceived to give voice to marginalized student populations traditionally underserved by public education in the presentation of objectively-assessed achievement gaps. As a result, one would expect to observe data use activities explicitly examining resources and supports targeted

towards the achievement of marginalized students, particularly in those schools regularly referring to student performance data. Surprisingly, while the underperformance of failing students and reclassification rates of English Language Learners was of vocal concern for many participants, few data use activities were described or observed in relation to the discussion of specific demographic subgroups.

Within this study, this may be in large part explained by the homogeneity of each schools' student population. Students from each school site were predominantly non-white and socio-economically disadvantaged; ELL students comprised about one-quarter of Belleworth's student body and nearly half of Woodson's. Conversations around performance improvement or targeted assistance were rarely observed to regard specific student sub-categories based on ethnicity, race, socio-economic status, or English language fluency and more likely to be held around the whole student body. The question underpinning schools' data use activities seemed to be: *how do we ensure that all of our students are learning well?* While school practitioners may well be thinking about how different groups are doing in relation to this question, this was not frequently observed as an explicit discussion.

Nevertheless, the success and achievement of individual students was a prominent concern for most teacher participants. Teachers were more likely to volunteer examples of progress based on individual students or the experience of their specific classes (i.e., Vayas, Kinsey, or "my first period"), rather than more cross-cutting categories of students (i.e., the girls or boys within this school), raising yet another tension with respect to the consideration of student performance data. An emphasis on understanding student and school performance through data entails the detection of trends and patterns of achievement within those data. That

is, the review of performance data intuitively directs one's attention to groups of students as they fall within certain bands of performance; and the units of analysis become subsections of the student body. At issue with at least some teacher participants, however, is the notion that the preeminence of data can also be experienced as a departure from the acknowledgment of individual student needs. That is, data can steer the conversation away from individuals, and in the extreme, school data conversations can distract teachers from getting to know their students on a personal level. At least one teacher noted that to use data to monitor and guide individual student performance – as a personal trainer might note a person's speed or angle while she runs on a treadmill in the determination of whether or not she is “pushing herself” – would require an unsustainable level of attention and resources given his current teaching context. The use of data to identify areas of subgroup performance, therefore, is more practically feasible than its diagnostic use for every student. Still, at the end of the day, there appeared to be a number of teachers debating the practical value of data if “all these numbers” do not contribute to a better understanding of the motivations and lives of their students.

### **Treating Classrooms as Laboratories**

Advocacy for the use of data to inform instruction at the classroom level, and the prospect that teachers can and should facilitate data-driven inquiry as a matter of good practice, imply the treatment of classrooms as experimental laboratories. Within this space, the teacher as experimental scientist collects empirical data on targeted pedagogies and adjusts her or his approach in view of those data. Continual improvement is seen to be the outcome of reiterative cycles of small-scale study and responsive modification. Many teacher participants recognized

the value of methodical instructional refinement, as well as the benefits of more analytical strategies to evaluating one's practice. A research perspective brings a new, sometimes substantial level of objective insight into the classroom and a disciplined approach to inquiry (Bryk, Gomez, & Grunow, 2011).

This advantage, however, is balanced by the practical demands of classroom teaching. Variation introduced into the laboratory setting by different students, the wide range and depth of their individual needs, and the varying learning character of each class as a collaboration of students, as well as the myriad job responsibilities and demands on teachers outside of strict instruction, inevitably influence experimental conditions. In fact, they are intrinsically part of the experimental setting. It has certainly been acknowledged that efforts in educational research and development have not aligned well with the real needs of schools, such that sustained and coordinated problem-solving in education has not yet been realized (Bryk & Gomez, 2008). This study contributes to this conversation by highlighting how efforts to study teacher practice and its effects on student learning must therefore honor the practical demands of classroom-based teaching and learning.

In our approach to engaging school stakeholders in classroom-based research guided by methods-based protocols, it must be remembered that data collection and analysis activities peripheral to instruction, such as those that entail labor intensive methods or which require in-depth technical capacity building, are bound to be regarded by teachers as extraneous, onerous, and not worth the effort. The minute they lose instructional focus, data use becomes irrelevant. Further still, this study has shown that when research ideals have taken precedence over teaching and learning objectives, negative impacts on instruction can result (such as in the selection of

departmental learning objectives based on what can be measured as opposed to what is of highest instructional priority). Even where teachers are encouraged to focus on data collection that is meaningful to their instruction and purposeful in its reach, it must be recognized that there is a necessary process teachers must experience through which specific data use methods are identified as efficient or practically effective. Data use routines that make sense in classroom contexts will vary across teachers, and those teachers must be given the time and space to vet and examine new approaches to classroom-based research.

Additionally, there is still room for additional research on what constitutes “rigorous” data in the context of classroom teaching, as well as what methods of data collection are aligned with both instruction and the criteria for rigor. Are there methodologically sound ways to observe student engagement over the course of a lesson, for example, without having to rely on a teacher’s dedicated attention to individual student tracking? Are teachers familiar with a sufficient variety of data collection methods and tools, and do they understand the methodological benefits and limitations of each approach?

Until we better understand what data use routines can be best streamlined within processes of teaching and learning, however, there remains an important boundary distinguishing research from instruction within schools. Where that line exists is culturally-dependent and unique to each site. As such, the promotion of data use in classrooms must entail the careful consideration of where instruction ends and research begins. Asking teachers to adopt the identity of researchers must be regarded as a system of capacity-building, requiring time and a protected space for trial and error.

## **Professional Development**

As has been emphasized throughout this study, data use within schools for purposes of decision-making is not a learned skill; it is a paradigmatic shift in the way schools operate. The development of schools' capacity to manage, facilitate, and participate in data use processes, then, is not just the work of short-term professional development or discreet data use interventions. Indeed, even with the involvement of schools in programs specifically targeted toward data use, past studies have found that school stakeholders lack not only data analysis skills (such as in the interpretation of student test scores), but also the ability to identify solutions and next steps in addressing diagnosed problems (Marsh et al., 2006; Timperley, 2009).

In the case of Woodson College Prep, coaching has been instrumental in guiding teachers through data use processes, providing contextually-appropriate insight and objective perspective, as well as helping teachers maintain a healthy sense of external accountability to data use routines. However, in-class coaching is sometimes viewed as an extravagant resource and perhaps unnecessary for seemingly routine processes such as data collection and analysis. Indeed, some teachers at Woodson expressed concern that their access to PDSA coaches was an unsustainable benefit, that they should perhaps forego coaching (despite its value) in order to mimic a more typical teacher experience of using PDSA. But, as a mentoring professor commented, "It's not unreasonable for any professional to have a coach. If Tiger Woods can have a coach, so can I." The field of education is familiar with instructional coaches and coaching for other types of professional development. If the goal is to assist schools in using data well, and using data effectively for instruction, long-term coaching for teachers in data use should be seriously considered.

Previous research suggests that the presence of data “experts” in schools promotes effective data use by providing teachers with support in identifying pertinent research, assisting them in managing and analyzing student data, and applying knowledge they gain from student data in making instructional decisions (Colbert & Kulikowich, 2006; Kerr, Marsh, Ikemoto, Darilek, & Barney, 2006; Quartz, Kawasaki, Sotelo, & Merino, 2014; Rock & Wilson, 2005). As the experience of Woodson’s teachers has shown, coaches have also been key in their provision of an “outside eye” to one’s personal practice, alongside their guidance of teachers through new processes of data collection, study, and interpretation. Additionally, the presence of coaches introduced an added layer of low-level accountability to new data use routines, providing external incentives to compile, study, and respond to data with timeliness and consistency.

But the presence of an expert does not in itself result in data use. Woodson’s own research professor, Dr. Baher, expressed some frustration with an overreliance on her expertise to conduct data use processes on behalf of faculty. Her university colleague cautioned Woodson’s teachers: “It’s not just about having a coach, but making a coach work.” That is, the ultimate point of expert involvement is to listen to practitioners and encourage data use practices in ways that enable instructional decision-making. In practice, this suggests that coaching and mentorship is most effective when offered alongside the actual involvement of teachers and administrators in processes of data identification, collection, analysis, and interpretation. Only through practice do these processes become integral into the “natural rhythms” of teachers and administrators. Regularly participating in data use activities offers the opportunity to gradually gain experiential knowledge in working with data. It also pushes teachers and administrators into needed dialogue concerning the alignment of teaching and learning objectives, data collection

instruments, measures of achievement or progress, and their implied meaning. It forces reflection on the questions: What would I like to know that would enable me to teach more effectively? How would I go about systematically measuring that? What are the implications of data collection on my practice? What tweaks in the system do we need to support data collection, opportunities to digest and understand those data, and then figure out how findings might lead to pedagogical or curricular changes? And, What technical skills do I need to build to ensure data use processes are meaningful for my practice? Active, guided participation in data use processes contributes to users' increased fluency in the language of data, and in turn, their perceived contributions to conversations using data.

### **Study Limitations**

In recognition of the limitations of data, this study is also subject to its own limitations as a body of research. In particular, the case study comparison – and perhaps all qualitative research – is subject to researcher bias in processes of data collection and analysis. Bias is discussed in several different ways but is sometimes referred to as a perspective which, in the positive, “reveals important aspects of phenomena that are hidden from other perspectives,” or in the negative, “is a perspective which obscures more than it reveals” (Khilnani, 1993). More neutrally, Becker (1966) argues that influential argument integral to sociological analysis “is always from someone’s point of view, and is therefore partisan.” The approach to this research has included multiple methods and data sources as a way of balancing perspective, invested substantial time at each school site and with participants to obtain saturation, and has solicited participant feedback as a way of “checking” research findings. Nevertheless, the positionality of the researcher, particularly one who identifies as an evaluator, is a persistent factor in how the

data are analyzed and interpreted. In the spirit of full disclosure, my own interest in this study topic stems from many years of work in educational settings, including as a classroom teacher, after-school tutor, curriculum developer, and program coordinator, as well as intensive work in education program evaluation. The latter persona reflects my own personal value for data and my own advocacy for data as a benefit to school programs and school stakeholders. This is balanced, however, by my experiential understanding of the demands of school administration and classroom instruction. The approach to this study was an honest attempt to challenge my own assumptions as to whether and how data are used within schools. By invoking the perspectives and experiences of practitioners, I hope to have done so with veracity.

The experiences of participants have been provided in rich detail with the intention that readers might identify with or relate to their perspectives. The method employed aims for transferability of study findings to school settings of interest as opposed to broad generalizability. A single case drawn from a purposeful sample is, by definition, not representative. As such, the views of teachers and principals expressed throughout this study belong only to those participants and are not intended to stand as voices for their colleagues or for their schools. They represent perspectives captured one moment in time that are subject to change and evolution with added experience.

## **Conclusion**

In response to the policy edict that schools should be using data for decision-making, this comparative case study investigates how school-based data are identified, prioritized, analyzed, interpreted, and used in schools. This inquiry has led to the exploration of a wide variety of perspectives surrounding schools' conventional use of data. These range from a complete distrust

of, and frustration with, the heightened public focus on data to the optimistic advocacy of data as a way of identifying practical points of leverage in improving student performance. Although individual perspective will always be a varying factor, this study has also found that the paradigms, contexts, and cultures characterizing each school are largely influential on how the value and utility of data are regarded by teachers and principals. That is, how decisions are made within a school, who leads and contributes to decision-making processes, what systems of data collection, composition, and review are in place within a school, what stakeholders identify as credible data, and how schools engage in data use activities and processes all factor into whether data are eventually put to use, or are not. Study findings thus suggest that the use of data in school-based decision-making does not follow a rationalist pattern guided by economic cognitive processing at the individual level or formal, normative structures at the institutional level.

“Decision-making” is considered the primary focus of data use in schools wherein teachers and principals make instructional and administrative choices informed by empirical evidence. This study shows, however, that because decision-making occurs at so many different levels within educational systems (i.e., classrooms, schools, and the District), and involves such a great variety of stakeholders (i.e., students, parents, teachers, principals, and District administrators), decision-making is not a singular activity but one that encompasses a number of different, and sometimes competing, purposes. Data are not only used to assess student achievement and gauge teacher effectiveness, but they are also used to prioritize areas for the investment of school resources and to develop classroom-based and school-wide instructional strategies. They are used as evidence to justify and modify pedagogical and organizational innovations, as well as to demonstrate compliance with District standards and policies. Data are

considered the substance of school accountability to the general public. Integrated into all of these activities, data have become a core component of school functioning. Purpose, as well as situational context, drives how and to what extent data are used in schools.

Indeed, with the proliferation of technological infrastructure and an ever-increasing public interest in monitoring and evaluating school, teacher, and student performance with data, schools find themselves in the midst of an overwhelming variety of options with which to gauge various aspects of student achievement and school success. This study has also explored how principals and teachers make sense of these data, how these data align with data collected in the course of their everyday work with students, and how school stakeholders determine what additional data should be considered to better inform instructional and organizational improvement. It is found that what principals and teachers regard as “credible” data relies on the purpose for which data are intended, the context in which they are collected, and the complexity of issues under investigation. As such, the interpretation of what data are credible within schools is more context-driven than criteria-driven.

Several key findings and learned lessons resulting from this study have been discussed as a way of promoting effective data use in schools. While data infrastructure and systems of data collection and review are certainly considered important in facilitating data use, it is found that they are not essential to data use in schools. Schools are seen to incorporate a host of different data into their day-to-day decisions, which may or may not rely upon routinized data use systems. On the other hand, transparent processes of decision-making and the authentic engagement of school stakeholders in decision-making are prerequisite to data use. Who determine what should be done with school data, and the extent to which stakeholders feel they

have control over decision-making processes, substantially influences whether and what data are actually referenced in making decisions. Dynamics of power and authority also come into play in distinguishing user “buy-in” to data use processes from a sense of proprietary ownership over data itself. There exists a careful balance between teachers’ perceived autonomy over the use of data in their schools and the establishment of a culture of mutual accountability among school stakeholders. These relationships are complicated by the variety of political and practical purposes motivating data use in schools which have been seen to pull teachers, principals, and district administrators in competing directions. Finally, alongside the recognition that what is valued as “credible” data may not fall within the bounds of what is commonly regarded as “rigorous, systematically-collected” data (such as anecdotal data), is the notion that data use within schools is dependent on the alignment of data, data collection processes, and expectations for use with instruction and instructional needs.

Importantly, what constitutes “effective data use in schools” is treated as the systematic integration of data into dialogue, deliberation, and decision-making supportive of teaching and learning. Critical to this perspective is the notion that data are a tool in understanding school effectiveness. Data are not a replacement for sound systems of strategic thinking, acute professional judgment, or established processes of inclusive decision-making. Data are not an outcome in and of themselves. The objective is not to move the data, nor to allow data to exclusively determine what we do in classrooms and in schools. Rather, the goal is to incorporate data into our understanding of what makes or could make for better schooling. Data are a means to evaluate where schools are and where they need to go. Use of data in this way is, therefore,

necessarily dependent on school culture and context – who schools serve, how these people are best reached, and the ways in which schools poise themselves to respond.

While none of the study findings are intended to be prescriptive, they have shed light on areas where processes of data use can be better suited for practical implementation in schools and classrooms. The study has found that “best practices” in data use cannot simply be introduced to principals and teachers as a way of producing better data use. Rather, school-level stakeholders must be regarded as the primary architects of data use processes, including the definition of what data should be collected and how. That is, in addition to the provision of resources (time, funding, capacity building, expert involvement, and the opportunity for collaboration), schools with strong systems of data use:

- 1. Start with identifying relevant questions of practice and building consensus around what sources of data reasonably speak to those questions.** Data do not speak for themselves. How data are intended to be used, and the ways in which they are collected, compiled, and analyzed, must be both transparent to and endorsed by teachers and principals. Conversations should also acknowledge and identify what data use practices individuals currently undertake to inform their understanding of an issue and how systematic or supplementary data collection would support these ongoing efforts.
- 2. Meaningfully involve stakeholders in the development and facilitation of measures and tools to collect those data.** Data collection must make sense in the context of instruction. The benefits and limitations of various methodologies should be transparently conveyed to data users. This holds true even when validated instruments are adopted or adapted, especially as “expert involvement” does not necessarily translate into stakeholder valuations of credibility.
- 3. Collaboratively review results among faculty in a dialogue about what the data do and do not say and why, adjusting data collection instruments and processes to fit instructional needs and demands in iterative rounds of implementation.** The cultivation of schools as models of continual improvement and classrooms as research labs will require a great deal of long-term technical capacity building on the part of teachers and administrators. The development of technical research skills through hands-

on practice is perceived to contribute to both an enhanced fluency with data use methods, as well as an increased level of trust in data use processes and applications. Importantly, efforts to create researchers out of school practitioners will need to be weighed against the need of teachers and principals to focus on their primary professional functions of instruction and school administration.

4. **Openly acknowledge competing influences on data use.** Data used for purposes of formative organizational development hold different meaning, and present different consequences from data that are meant to serve interests of public accountability. These competing purposes not only change the interpretation of data, but can also affect processes of data collection, compilation, and instruction. Similarly, expectations that classroom-collected data will be used not only to inform instructional development but also to assess school performance influence teacher perceptions of ownership and control over data. Forthright conversations about the interests, motivations, and expectations guiding data use might surface otherwise covert misgivings about data use processes.

This research extends our knowledge of the context surrounding data use in schools by looking in detail at the ways school practitioners engage and interact with data in the course of their everyday work. It recognizes that data are only one aspect of “good practice” demanding the attention of teachers, principals, and district administrators and, therefore, that data use is subject to individuals’ interpretive processes of noticing data, prioritizing “credible” data, making sense of those data in view of their perceived purposes, and finding practical avenues of feeding data into decision-making processes. Importantly, this study contributes to our knowledge base by examining how teachers apply a wide variety of data in instructional activity, as well as the positive and negative influences of data use on pedagogy, curricular strategy, and student assessment. It also details the many ways that administrators use data to understand how well their schools are doing, and how this perspective interacts with teachers’ sometimes disparate regard to data. Findings show that organizational cultures influence processes of individual and whole-school data use, particularly with respect to systemic issues of decision-

making authority, teacher and school autonomy, and incentives and disincentives imposed by implicit and explicit expectations of accountability. In showcasing the many specific ways that school stakeholders are both confronted by and involved with data, this research contributes to our understanding of data use as a nuanced, contextually seated endeavor. Using data effectively is not simply a matter of doing nor is it a direct outcome of resource investment. Rather, this study has shown that data use occurs within schools in a variety of different ways, for a variety of different motivations, and with a variety of different results.

The terrain surrounding effective data use in schools is vast and challenging. As school stakeholders continue to gain more experience and facility with data, so will our understanding of what does and does not work in various circumstances. Until then, while the ways in which data are used in schools may be conditional, the prevalence of data is not. It is hoped that this study contributes to an enhanced understanding of what this means within schools and among their many stakeholders.

## Appendix A

### Case Study Coding Framework

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>Accountability <ul style="list-style-type: none"> <li>“Dog and pony show”</li> <li>Stakeholder expectations</li> <li>Attraction to Pilot School</li> </ul> </li> <li>Becoming a Pilot School <ul style="list-style-type: none"> <li>Building student programming</li> <li>Concept to action</li> <li>New school challenges</li> <li>Pilot School learning curve</li> </ul> </li> <li>Community <ul style="list-style-type: none"> <li>Communication with parents</li> <li>Investment in Pilot Schools</li> <li>Parent participation</li> </ul> </li> <li>Data - Not valuable <ul style="list-style-type: none"> <li>Data vs. information for instruction</li> <li>Focus on numbers</li> </ul> </li> <li>Data - Valuable <ul style="list-style-type: none"> <li>Data for instruction</li> <li>Meaningful assessment</li> <li>Meaningful feedback</li> <li>School-based measures</li> </ul> </li> <li>Data Analysis <ul style="list-style-type: none"> <li>Framing</li> </ul> </li> <li>Data Collection <ul style="list-style-type: none"> <li>Classroom environment</li> <li>Identifying measures</li> <li>Interpreting standards</li> <li>Teacher observation</li> </ul> </li> <li>Data Culture <ul style="list-style-type: none"> <li>Assessment autonomy</li> <li>Building relationships</li> <li>Data driven cycles <ul style="list-style-type: none"> <li>Natural rhythms</li> <li>Practicing processes</li> </ul> </li> <li>Dedicating resources</li> <li>Goal development <ul style="list-style-type: none"> <li>Focus on student achievement</li> </ul> </li> <li>Internal accountability <ul style="list-style-type: none"> <li>Defensiveness</li> <li>Ownership &amp; individual autonomy</li> </ul> </li> <li>“Our story” <ul style="list-style-type: none"> <li>“It is what it is”</li> </ul> </li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>Personal propensity towards data</li> <li>Persons responsible</li> <li>Primary questions</li> <li>Understanding the purpose of data</li> <li>Data Distribution <ul style="list-style-type: none"> <li>Turnaround time</li> </ul> </li> <li>Data Interpretation <ul style="list-style-type: none"> <li>Demographics</li> <li>Longitudinal data</li> <li>Reliability</li> <li>System complexity</li> <li>Year-to-year differences</li> </ul> </li> <li>Data Sources <ul style="list-style-type: none"> <li>Anecdotal information</li> <li>Attendance</li> <li>Behavioral vs. academic</li> <li>High School Exit Exam</li> <li>Classroom observations</li> <li>College acceptance rates</li> <li>Cumulative files</li> <li>English Language test data</li> <li>Enrollment</li> <li>Formative assessment <ul style="list-style-type: none"> <li>Hidden curriculum in assessment</li> </ul> </li> <li>Scoring <ul style="list-style-type: none"> <li>Test development</li> </ul> </li> <li>Gifted &amp; Talented Programs (GATE)</li> <li>Grading <ul style="list-style-type: none"> <li>Grade check</li> </ul> </li> <li>Graduation</li> <li>Individual Education Plans (IEP)</li> <li>Informal Observations <ul style="list-style-type: none"> <li>Student background</li> <li>Student report</li> <li>Teacher reputation</li> <li>Teacher self-report</li> </ul> </li> <li>Lack of data <ul style="list-style-type: none"> <li>Lack of standardized state exams</li> </ul> </li> <li>Learning management systems</li> <li>Parent surveys</li> <li>Personal experience</li> <li>Pilot School Review</li> <li>Program description</li> </ul> </li></ul> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>Standardized testing <ul style="list-style-type: none"> <li>Common Core</li> </ul> </li> <li>Student discipline</li> <li>Student survey</li> <li>Student work</li> <li>Teacher survey</li> <li>Teacher-led evaluation</li> <li>Title I</li> <li>Value-Added Models</li> <li>What isn't measured</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | <ul style="list-style-type: none"> <li>Pilot School as a reform model <ul style="list-style-type: none"> <li>Innovation</li> <li>Pilot School power</li> <li>Pilot School strategy</li> <li>Relationship with comprehensive schools</li> <li>Small learning communities</li> <li>Small schools</li> <li>Pilot Schools vs. Comprehensive schools</li> </ul> </li> <li>United Teachers Los Angeles (UTLA)</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <ul style="list-style-type: none"> <li>Data Use <ul style="list-style-type: none"> <li>Budget decisions</li> <li>District data use</li> <li>Incorporating data into practice</li> <li>Misuse <ul style="list-style-type: none"> <li>Manipulation</li> <li>Punitive</li> </ul> </li> <li>Motivating</li> <li>Multiple purposes</li> <li>Overwhelming data</li> <li>Parent data use</li> <li>Privacy</li> <li>Program monitoring</li> <li>Research-based evidence</li> <li>Reviewing results (or not)</li> <li>Self-promotion</li> <li>Student data use</li> <li>Teacher self-reflection</li> <li>Technical capacity</li> <li>Unsystematic</li> </ul> </li> <li>District Political Context <ul style="list-style-type: none"> <li>"Adult issues"</li> <li>Bureaucracy <ul style="list-style-type: none"> <li>District "conditioning"</li> </ul> </li> <li>Conspiracy <ul style="list-style-type: none"> <li>Privatizing education</li> <li>Setting schools up to fail</li> </ul> </li> <li>District - Perceived lack of support</li> <li>District - Perceived support</li> <li>Market competition <ul style="list-style-type: none"> <li>Pilot School innovations</li> <li>School choice</li> </ul> </li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>Evaluation Activities <ul style="list-style-type: none"> <li>High stakes evaluation</li> <li>Pressure</li> <li>Student stress</li> <li>Teacher job security</li> </ul> </li> <li>Continuous improvement</li> <li>No reference to data</li> <li>Plan-Do-Study-Act (PDSA) <ul style="list-style-type: none"> <li>PDSA effectiveness</li> </ul> </li> <li>Peer teacher mentoring</li> <li>Pilot School Review</li> <li>Principal Review</li> <li>Response to Intervention</li> <li>Single Plan</li> <li>Student panels</li> <li>Teacher evaluation</li> <li>Western Association of Schools and Colleges (WASC)</li> <li>Good Teachers <ul style="list-style-type: none"> <li>Perfecting the practice</li> <li>Performance incentives</li> <li>Personal connections with students</li> <li>Professional development</li> <li>Professional judgment</li> <li>Pulling weight</li> </ul> </li> <li>Identifying student need <ul style="list-style-type: none"> <li>Failing students</li> <li>Lost in the crowd</li> </ul> </li> <li>My Integrated Student Information System (MiSiS)</li> </ul> |

#### Organizational change

- Challenges
- Collaboration
  - Buy-in
  - Common language
  - Consistency & alignment
- Decision-making
  - Competing interests
  - Decision makers
  - Non-transparent decision-making
- Leadership
  - Changes in administration
- Pilot School autonomy
- Sustainable systems
- TrustOutcomes based instruction
- Lesson planning

#### Pilot School Instruction

- Additional supports to students
- AP Courses
- Bell schedule
- English Language Development (ELD)
- Full inclusion
- Inter-Disciplinary Projects (IDP)
- Technology

#### Pilot School Operations

- "There are a lot of moving parts"
- Counseling
- English Learning Advisory Committee (ELAC)
- Election-to-Work Agreement (EWA)
- Facilities
- Governing School Council (GSC)
- Instructional Learning Team (ILT)
- Budget & funding
- Multiple hats
- Pilot school challenges
- Personnel recruitment
- School-based policy implementation
  - Technology
  - Time
- School Culture
  - Pedagogical philosophies
  - Pilot population
  - Pilot school "fit"
  - Teacher driven
  - Teacher support
  - Teachers vs. Administration
  - Valuing curricular content
  - Valuing parents
  - Valuing students
  - Valuing Teachers
- Understanding Student Progress
  - Career readiness
    - Student skill building
  - Social promotion
  - Socio-emotional
    - Character education
  - Student motivation
  - Student orientation

## **Appendix B**

### **Guiding Questions for School Leaders in Supporting the Effective Use of Data in Decision-Making**

#### **Understanding the Influence of Data Needs and Purposes**

1. What is the intended purpose (i.e. Instructional change, organizational learning, or performance and accountability) for each of my school's data activities?
2. How are my school's data and data-based activities expected to contribute in response to these needs and purposes?
3. How does each purpose influence what data should be collected, how that data should be collected, and in what ways data should be analyzed and interpreted?
4. Do my schools' stakeholders understand and endorse the purpose of each data activity?
5. Are certain data sources or data collection activities expected to respond to multiple purposes? If so:
  - a. How, if at all, does the analysis and interpretation of the data change when it is intended to serve another purpose?
  - b. How might these differences affect the ways in which the data are ideally collected and compiled?
  - c. How might these differences impact stakeholders' understandings of, and value for the data?

#### **Understanding the Influence of Stakeholder Perspectives**

1. What do the processes of data collection, analysis, interpretation, and dissemination look like from the perspective of each stakeholder (group)?
2. How do stakeholders' pre-conceptions of the value and utility of data, their perceptions of how data will be used, and their own personal capacity in using data influence my school's collective approach to data use?
3. How might different perspectives on teaching, learning, and/or school administration held by different stakeholders influence the ways in which they view and understand my school's use of data?
4. How do beliefs, motivations, and knowledge about data use vary both within and between stakeholder groups?

#### **Understanding the Influence of Decision-Makers and Decision-Making Processes**

1. Who is (officially or unofficially) designated to make what decisions within my school?
2. Do school stakeholders endorse these decision-makers? What are the power relations/dynamics between these groups?

3. How are decisions made? Are these decision-making procedures transparent? Are they routinely implemented? Are school stakeholders committed to these decision-making procedures?
4. Has my school established a sufficient level of trust amongst school faculty, administration and staff to engage in constructive conversations about our teaching and learning practices (irrespective as to whether these conversations involve looking at data)?
5. Do stakeholders feel that their voices are valued in decision-making processes and that they consistently bear weight when decisions are made?

### **Thinking About Data Systems and Structures**

1. Does my school have the infrastructure to access, extract and compile the data we need in a timely way? Are the data accurate?
2. Have we established systems to regularly collect, organize, process, and review data?
3. Who is responsible for carrying out these responsibilities?
4. Are adequate resources (time, personnel, funds) in place to support my school's approach to data collection, data analysis, the interpretation of results, and reporting?
5. Are the types of data we collect responsive to our data needs and purposes? Do the data respond to the needs of my school's decision-makers? (i.e., Have we established a demand for systems of data review?)

### **Identifying Credible Data**

1. What are the data that each of my school's stakeholders value as credible, and why?
2. Are there differences amongst individuals or stakeholder groups in the types of data they value most? If so, is this a reflection of their value for the data themselves, the intended use/interpretation of the data, or something else?
3. How are perspectives of credibility influenced by the explicitly and tacitly communicated purposes of the data?
4. How are perspectives of credibility influenced by what my school's stakeholders know, or do not know, about the methods used to collect and analyze data?
5. What efforts have we made as a school to collectively identify our goals and objectives, and to align these with the types of activities we implement and the kind of data we collect on those activities?
6. Are the data we incorporate into processes of decision-making relevant to the questions we have as a school? Do they provide meaningful information to our collective consideration of those questions?

### **Thinking About My School's Data Use Processes**

1. Do my school stakeholders have the technical capacity to analyze and understand the data we have?

- a. To what extent have my school stakeholders been involved in the collection, compilation, analysis, and review of the data? How might their degree of involvement in these activities impact their ability to interpret and make use of the results?
2. When we review data, do we articulate our beliefs, motivations, and knowledge surrounding the topic of focus? For example, do we discuss:
  - a. What we expect to see in the data, and then, whether the results match our expectations? If we observe any differences, do we discuss why this might be so?
  - b. What we know about how data are collected and how they were analyzed?
  - c. What we think about what the data do, or do not represent?
  - d. Whether the data, even if limited, provide us with useful insight?
3. Are we able to collectively strategize ways in which we might be able to improve our performance based on what we observe from the data in service to our goals and objectives?
  - a. In reviewing our data, what are some immediate next steps teachers can take to translate student needs implied by the data into instructional change?
4. Are there ways that we can improve the measurement of our activities to ensure the resulting data are meaningful to our practice? Have we considered the tradeoffs between modifying our measurements and maintaining our measurements so that we can consistently track our performance from year-to-year?
5. How, and how well are we communicating our data with stakeholders? How might this influence the way we use data within our school?

## **Appendix C**

### **Teacher Interview Protocol (Semi-Structured)**

*This semi-structured interview protocol contains a list of possible questions to be drawn from in interviews with pilot school teachers. As a flexible framework, questions may be added or omitted from the interview in response to participant feedback. Topics of discussion, however, are expected to stay within the content areas detailed below for purposes of research.*

#### **Interview #1**

##### *Understanding teaching objectives*

1. Tell me about how you came to teach at this school.
  - a. Was there anything particularly intriguing to you about working here?
2. As a teacher, what would make this a “successful” year for you and your students?
3. As a pilot school, are there characteristics of the school that especially support your success as a teacher?
  - a. How do you see these school characteristics applied in practice? Examples?
4. In your experience, do the school’s exercise of autonomies support your success as a teacher?
  - a. Can you provide an example of how you have experienced this connection, or;
  - b. Can you provide an example of how you anticipate experiencing this connection?

##### *Perceptions of “information”*

5. Tell me about how you gauge student learning in your classroom. (Walk through several examples)
  - a. How do you know if/when a student is in need of extra supports? (Probe for “every day” practices as well as more formal assessments)
  - b. Do you find some methods of assessment more useful than others?
  - c. Do you know if the way you assess student learning is similar to the practices of other teachers in this school?
6. Would you say that your students know if they are doing well in your class (or not)?
  - a. How would they know this?
7. If a student does receive extra supports from you or other school programs, how do you know these are helping?
8. If you were interested in understanding how well your class was doing in relation to other X Grade [subject] classes, what kinds of information would you look to? Why?
  - a. Is this something you have done? Why or why not?
9. If someone were to ask you how well your school was doing. How would you respond?
  - a. What kind of information might you offer them to support your case? Why?
  - b. What kinds of information do you think your school stakeholders (parents, support organizations, principal, district personnel, etc.) find most convincing? Why?

### *Understanding accountability requirements*

10. What kinds of data are you expected to collect throughout the school year?
11. What kinds of information might you collect throughout the school year that are not required?
  - a. Do you know if other teachers do the same?
12. How, if at all, are these data collection activities different from last year, or previous years?
  13. In your opinion, has meeting these accountability requirements contributed to your success as a teacher? To your school's success?

## **Interview #2**

### *Perceptions of data use*

1. There appears to be a trend, where “data used for decision-making” is considered a best practice in schools. Have you heard of this phrase? In your opinion, what does it mean?
2. In your opinion, how (if at all) are these practices actually applied, particularly in the classroom?
  - a. Are you familiar with the kinds of data your school is meant to regularly collect?
  - b. Are you expected to collect any kinds of “data” in your classroom? Examples?
  - c. Are you expected to interpret either school “data” or “data” used in your classroom? Examples?
  - d. Are you expected to incorporate any types of “data” into your teaching? For what purposes? Examples?
3. Do you find yourself conducting any of these activities?
  - a. If so, can you walk me through a few examples of what this looks like?
  - b. If not, why not?
4. Are the kinds of data that you might collect to inform your own teaching different from the types of data your school uses to present its performance? Examples?
5. In your opinion, what kind of information is most useful to you in improving what your students learn?
6. In your opinion, what kind of information is most useful to you in improving your success as a teacher?
7. We have talked about data that you collect for use in your own classroom, as well as data that the school collects to report on its own performance for purposes of accountability. Are these types of information different to you? If so, how?
  - a. Do you use these types of data differently? If so, how?
  - b. Is one type of data more useful than the other?

### *Capacity and Culture*

8. How would you characterize your own level of comfort working with “data?”
  - a. How would you characterize your interest in doing so?
9. How would you characterize other teachers’ level of comfort working with “data?”
  - a. How would you characterize their interest in doing so?
10. How would you characterize your principal’s level of comfort working with “data?”
  - a. How would you characterize his/her interest in doing so?

11. Do you find yourself participating in discussion involving data with your colleagues?  
Examples?
12. Does your school distribute or present data reports to students and parents?
  - a. How would you characterize their level of comfort understanding this type of “data?”
  - b. How would you characterize your parents’ interest in school data?
13. What skills do you believe you need to have in order to interpret and use the information your school produces?
14. What areas of skill development do you think would be important for you to interpret and use the information your school produces?
15. Are there areas of teaching, learning, or other school activities that you wish you knew more about but for whatever reason, are unable to obtain richer information on?
  - a. Examples?
  - b. What kinds of information would you want to have?
  - c. What might prevent you from obtaining this information?

### **Interview #3**

#### *Perceptions of data use policies & tools*

1. In your opinion, does your school support the use of data and information to improve teaching and learning? Examples?
  - a. Do you feel supported by other teachers within the school to use data and information to improve teaching and learning? Examples?
  - b. Do you feel supported by your principal to use data and information to improve teaching and learning? Examples?
2. Do you feel that your school is successful in reporting on its own successes and challenges using data collected either at the school level or by you in your classroom?
  - a. What, in your opinion, contributes to this?
3. Do you feel that whatever policies and expectations exist at the District level support your use of data in your classroom? Your school’s use of school data?
4. Are there different policies, expectations, or incentives either at the District or school level that you think would better support your use of data in your classroom?
5. What, in your opinion, are the key elements of a school data system that collects, analyzes and reports information efficiently?
6. What, in your opinion, are the key elements of a school data system that make use of that information?
7. How much time would you estimate you personally spend fulfilling data requests?
  - a. Do you feel that this is reasonable alongside your other teaching responsibilities?
8. In your opinion, is the effort required to meet accountability requirements and data requests equivalent to the benefits you receive from this information? Examples of why/why not?
  - a. If not, what might a fairer balance look like?

#### **Interview #4:**

I would like to revisit some of the discussion points we had throughout the school year and get a sense of whether your opinion has changed at all since we last talked about them.

##### *Perceptions of “information”*

1. At the beginning of the year, I asked you, “If a potential parent were to ask you how well your school was doing. How would you respond?” You suggested that....
  - a. Given all of the activities that have taken place at your school this year, how might you answer this question now?
  - b. What kind of information might you offer them to support your case? Why?

##### *Perceptions of data “use”*

2. Over the course of the school year, you have engaged in several types of data collection activities [list]. In your opinion, is the effort required to collect and supply this information equivalent to the benefits you receive from this information? Examples of why/why not?
  - a. If not, what might a fairer balance look like?
3. Looking back on this year, what makes data difficult to use to improve student learning? What has supported either you or your school in using data to improve student learning?
4. Looking back on the year, what skills do you believe you, your colleagues, principal, and parents need to have in order to interpret and use the information your school produces? Example?
5. What areas of skill development do you think would be important for you in order to interpret and use the information your school produces in the year to come?

##### *Perceptions of data use policies and tools*

6. Are there district or school policies that you could identify as being instrumental in promoting/hindering the use of information in informing your teaching?
7. Earlier, I had asked you what might be the key elements of a school data system that collects, analyzes and reports information efficiently? You said....
  - b. Thinking back on this school year, do you have anything to add or amend to this characterization?
8. I also asked you what might be the key elements of a school data system that makes use of information?
  - c. Thinking back on this school year, do you have anything to add or amend to this characterization?

##### *Concluding remarks*

9. In your opinion, is using data to improve student learning a reasonable expectation? A worthwhile endeavor?
10. Are there any topics I have not addressed over the course of interviews together that you would like to raise?

## **Appendix D**

### **Principal Interview Protocol (Semi-Structured)**

*This semi-structured interview protocol contains a list of possible questions to be drawn from in interviews with pilot school principals. As a flexible framework, questions may be added or omitted from the interview in response to participant feedback. Interview questions, however, are expected to stay within the content areas detailed below for purposes of research.*

#### **Interview #1**

##### *Understanding school performance objectives*

1. Tell me about how your school became a pilot school.
2. As a pilot school, what autonomies is your school exercising to achieve its vision/mission?
  - a. What do these autonomies look like as they are practiced in your school?
3. Tell me about how you see the exercise of these autonomies as contributing to your school's ability to meet its vision/mission?
  - a. Can you provide an example of how you have seen this accomplished, or;
  - b. Can you provide an example of how you anticipate seeing this accomplished?

##### *Perceptions of "information"*

4. If a potential parent were to ask you how well your school was doing. How would you respond?
  - a. What kind of information might you offer them to support your case? Why?
  - b. Do you have discussions with your staff and faculty about the progress of your school?
    - i. What kinds of things do you talk about?
  - c. What kinds of information do you think your school stakeholders (parents, support organizations, teachers, district personnel, etc.) find most convincing in determining the health or success of your school? Why?
    - i. Are these different for different stakeholders?
5. As the principal, how do you know if something needs to be "fixed" or changed in the way your school approaches its teaching and learning activities? Examples?
6. Who else in your school might you identify as someone who contributes to the development of school policy or strategy?
7. I am very interested in the notion that "data used for decision-making" should be promoted as a best practice in schools. In your opinion, what does this practice entail?
  - a. From your perspective, how are such practices applied?
  - b. Can you help me understand how your school may be using data for decision-making purposes? [Think of different types of decision makers]

##### *Understanding school accountability requirements*

8. I understand that your school is expected to meet a number of different accountability requirements. Could you describe these to me in general?
  - a. Cross-check researcher knowledge of requirements based on document review

9. Is there a timeline or schedule for this year on which you are attempting to meet these requirements? Would you be able to describe this for me?
  - b. If not, which of these requirements do you expect to attend to this year?
10. How, if at all, are these activities different from last year, or previous years?
11. Can you tell me more about how you go about developing this timeline/plan? How do you prioritize the order in which you address your various accountability requirements?
  - a. Can you walk me through a time when you had to make a decision about which requirements you were going to meet first?
12. How do you intend to go about meeting your accountability requirements for this year?
  - a. Who does this involve?
  - b. What are the processes of data collection, input, analysis, and dissemination?
13. In your opinion, does the data you are expected to collect help you to evidence your achievement of your school's vision and mission? Examples?
  - a. If not, what might be better ways of showing your school's achievements and progress?

## **Interview #2**

### *Perceptions of data use*

1. In your view, who are the "end users" of the accountability data your school produces?
2. How do you envision these people making use of the data your school produces?  
Examples for each set of stakeholders identified?
3. Do you see yourself using the data your school collects for purposes of accountability to inform your own decision-making?
  - a. If so, how? Tell me about a time when you found this data useful.
  - b. If you do not, why not? Tell me about a time when you did not find this data useful.
4. Are there other types of information your school collects outside of required accountability data? Why/why not?
5. Do you see yourself using data (for accountability requirements or otherwise) in your day-to-day practice? Examples?
6. Do you see your teachers using data in their day-to-day practice? Examples?
7. Do you see your parents making use of the data your school collects? Examples?
8. In your opinion, what makes (or, what would make) data most useful in influencing your own management decisions?
9. In your opinion, what makes data difficult to use in making school management decisions?

### *Perceptions of culture and capacity*

10. Do you feel that your teachers are generally comfortable with using data to inform their teaching and learning activities? Examples?
11. Do you feel that your teachers are generally comfortable using data to determine the strengths and challenges of school programming other than their classroom activities? Examples?
12. What, in your opinion, would support teachers' comfort with using data in their classrooms?

- a. ...in determining the strengths and weaknesses of other school programming?
- 13. What skills do you believe you, your teachers, staff, and parents need to have in order to interpret and use the information your school produces?
- 14. What areas of skill development do you think would be important for you, your teachers, staff, and parents in order to interpret and use the information your school produces?
- 15. Are there areas of teaching, learning, or school management within your school that you wish you knew more about but for whatever reason, are unable to obtain richer information on?
  - a. Examples?
  - b. What kinds of information would you want to have?
  - c. What might prevent you from obtaining this information?

### **Interview #3**

#### *Perceptions of data use policy & tools*

- 1. Do you feel personally incentivized to use data in making management decisions about your school? How so/why not?
- 2. Are there district policies or incentives that you could identify as being instrumental in promoting/hindering the use of information in informing your management practices? Examples?
  - a. ...in informing teachers' classroom practices? Examples?
  - b. ...in informing parents of school progress? Examples?
- 2. Are there policies or incentives you have put in place that you believe are instrumental in promoting the use of information within your school? Examples?
- 3. What, in your opinion, are the key elements of a school data system that is successful in collecting, analyzing and reporting information?
- 4. What, in your opinion, are the key elements of a school data system that ends up using that information in practice?
- 5. Would you identify any key elements that would be detrimental to a school data system in collecting, analyzing and reporting information?
  - a. ...in actually using that information in practice?
- 6. How much time would you estimate you personally spend per week/month fulfilling data requests?
  - a. Do you feel that this is manageable alongside your other responsibilities as principal?
- 7. What other staff participate in fulfilling data requests? How time per week/month do you estimate they each spend on these activities?
  - a. Do you feel that this is reasonable alongside other responsibilities your staff attend to?
- 8. In your opinion, is the effort required to meet accountability requirements and data requests equivalent to the benefits you receive from this information? Examples of why/why not?
  - b. If not, what might a fairer balance look like?

#### **Interview #4:**

I would like to revisit some of the discussion points we had throughout the school year and get a sense of whether your opinion has changed at all since we last talked about them.

##### *Perceptions of “information”*

1. At the beginning of the year, I asked you, “If a potential parent were to ask you how well your school was doing. How would you respond?” You suggested that....
  - a. Given all of the activities that have taken place at your school this year, how might you answer this question now?
  - b. What kind of information might you offer them to support your case? Why?

##### *Perceptions of data “use”*

2. Earlier, I had asked you what factors would make data collected by your school most useful in your own management decisions? You said...
  - c. Reflecting on the kinds of data collected by your school this year, would you change your answer at all?
  - d. Would you say that data your school has collected this year represent these characteristics? Why or why not?
    - i. If yes, did you find yourself using this data for purposes of making school management decisions? Examples?
    - ii. If not, what do you think would have to happen in order for school-collected data to look like this?
3. In retrospect, what makes data difficult to use in making school management decisions?
4. Looking back on the year, what skills do you believe you, your teachers, staff, and parents need to have in order to interpret and use the information your school produces? Example?
5. What areas of skill development do you think would be important for you, your teachers, staff, and parents in order to interpret and use the information your school produces in the year to come?

##### *Perceptions of data use policies and tools*

6. Are there district policies that you could identify as being instrumental in promoting/hindering the use of information in informing your management practices this year?
  - e. ...in informing teachers’ classroom practices?
  - f. ...in informing parents of school progress?
7. Are there policies or incentives you have put in place that you believe are instrumental in promoting the use of information in informing school practices this year? Examples?
8. Earlier, I had asked you what might be the key elements of a school data system that is successful at collecting, analyzing and reporting information? You said....
  - g. Thinking back on this school year, do you have anything to add or amend to this characterization?
9. I also asked you what might be the key elements of a school data system that makes use of information? You said...

- h. Thinking back on this school year, do you have anything to add or amend to this characterization?
- 10. Over the course of the school year, you have engaged in several types of data collection activities [list]. In your opinion, is the effort required to collect and supply this information equivalent to the benefits you receive from this information? Examples of why/why not?
  - i. If not, what might a fairer balance look like?

*Concluding remarks*

- 11. As more schools continue to apply for pilot school status, what advice might you give for new schools entrants about how to navigate accountability and evaluation requirements?
- 12. Are there any topics I have not addressed over the course of interviews together that you would like to raise?

## **Appendix E**

### **District Personnel Interview Protocol (Semi-Structured)**

*This semi-structured interview protocol contains a list of possible questions to be drawn from in interviews with district personnel overseeing pilot schools. As a flexible framework, questions may be added or omitted from the interview in response to participant feedback. Interview questions, however, are expected to stay within the content areas detailed below for purposes of research.*

#### *Personal Background*

1. I understand your work w/ PS began as a [insert job title]. Could you tell me more about that?
2. And then you moved on to becoming a [insert job title]. What did this entail?
3. And now you are in a new role – could you tell me about this?

#### *Understanding pilot school performance*

4. Could you help me understand how PS become PS? Walk me through this process.
5. It sounds like you have an intimate understanding, then, of PS coming from a number of different perspectives.
  - a. From a District perspective, what—in your opinion—are some of the expectations for Pilot Schools? (What is it hoped that they will achieve?) These might be formal or informal.
  - b. In your opinion and experience, how are these expectations interpreted within schools?
    - i. Do they coincide with school-based objectives? Clash?
6. In your own opinion, what makes a PS a successful PS?
  - a. How do you see schools making sense of how well they are doing?
  - b. How do you see the District making sense of how well they are doing?
7. Last time we talked a little bit about how PS performance is reviewed by the District. You mentioned that PS participate in an Annual Performance Review. Can you walk me through this process?
  - a. Different for every school? In what way?
  - b. Different depending on Instructional Director? What was your specific approach? How might this be different from others'?
  - c. Not an evaluation – what is this process meant to achieve?
  - d. What is done with this information? Who looks at it?
8. External Team Review – Have you participated in any of these? Would you be able to walk me through this process?
  - a. What is this process meant to achieve?
  - b. What is done with this information? Who looks at it?
9. Aside from these activities, are PS expected to participate in any other performance reviews?

10. Last time you mentioned District was going through the exercise of determining criteria for maintaining PS status. Can you tell me about this process?
  - a. You also mentioned in our last meeting that you thought the PS model was one set up to support innovation rather than to set criteria and reward for meeting those criteria.
  - b. Can you tell me a little bit more about this approach? How do you see this taking effect in PS?
  - c. How might this approach have influenced your own work?

*Concluding remarks*

11. Anything else I haven't asked about that you would like to raise?

## References

- Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, 16, 3–9.
- Alkin, M. C. (2004). *Evaluation roots: Tracing theorists' views and influences*. Sage.
- Alkin, M. C., & Coyle, K. (1988). Thoughts on evaluation utilization, misutilization and non-utilization. *Studies in Educational Evaluation*, 14(3), 331–340.
- Allaire, Y., & Firsirotu, M. E. (1984). Theories of organizational culture. *Organization Studies*, 5(3), 193–226.
- Anckar, D. (2007). Selecting Cases in Cross-National Political Research. *International Journal of Social Research Methodology*, 10(1), 49–61.
- Bickel, W. E., & Cooley, W. W. (1985). Decision-oriented educational research in school districts: The role of dissemination processes. *Studies in Educational Evaluation*, 11(2), 183–203.
- Bryk, A. S., & Gomez, L. (2008). Reinventing a research and development capacity. *The Future of Educational Entrepreneurship: Possibilities for School Reform*, 181–206.
- Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting Ideas into Action: Building Networked Improvement Communities in Education. In T. M. Hallinan (Ed.), *Frontiers in Sociology of Education* (pp. 127–162). Dordrecht: Springer Netherlands.
- Campbell, D. T. (1975). “Degrees of Freedom” and the Case Study. *Comparative Political Studies*, 8(2), 178–193.
- Coburn, C. E. (2010). Partnership for district reform: The challenges of evidence use in a major urban district. *Research and Practice in Education: Building Alliances, Bridging the Divide*, 167–82.
- Coburn, C. E., Honig, M. I., & Stein, M. K. (2009). What’s the evidence on districts’ use of evidence. *The Role of Research in Educational Improvement*, 67–87.
- Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research & Perspective*, 9(4), 173–206.
- Coburn, C. E., & Turner, E. O. (2012). The practice of data use: An introduction. *American Journal of Education*, 118(2), 99–111.

- Colbert, R. D., & Kulikowich, J. M. (2006). School counselors as resource brokers: The case for including teacher efficacy in data-driven programs. *Professional School Counseling*, 9(3), 216–222.
- Court, J., & Young, J. (2003). *Bridging research and policy: Insights from 50 case studies*. Overseas development institute (ODI).
- Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56(3), 331–364.
- Cousins, J. B., & Leithwood, K. A. (1993). Enhancing knowledge utilization as a strategy for school improvement. *Science Communication*, 14(3), 305–333.
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., ... Weiner, S. S. (1985). *Toward reform of program evaluation*. JSTOR.
- Daly, A. J., & Finnigan, K. S. (2011). The ebb and flow of social network ties between district leaders under high-stakes accountability. *American Educational Research Journal*, 48(1), 39–79.
- Datnow, A., Park, V., & Wohlstetter, P. (2007). Achieving with data: How high-performing elementary systems use data to improve student achievement. *Los Angeles: Center on Educational Governance, University of Southern California*.
- David, J. L. (1981). Local Uses of Title I Evaluations. *Educational Evaluation and Policy Analysis*, 27–39.
- Dembosky, J. W., Pane, J. F., Barney, H., Christina, R., & Education, R. (2006). *Data driven decisionmaking in Southwestern Pennsylvania school districts*. RAND.
- Donaldson, S. I., Christie, C. A., & Mark, M. M. (2014). Credible and actionable evidence: The foundation for rigorous and influential evaluations. SAGE Publications.
- Erickson, F. (1986). Qualitative methods in research on teaching (pp. 119–62). Institute for Research on Teaching.
- Feldman, J., & Tung, R. (2001). *Whole school reform: How schools use the data-based inquiry and decision making process*. Citeseer.
- Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L., & Katzaroff, M. (1999). Mathematics performance assessment in the classroom: Effects on teacher planning and student problem solving. *American Educational Research Journal*, 36(3), 609–646.

- Gabbay, J., & May, A. le. (2004). Evidence based guidelines or collectively constructed “mindlines?” Ethnographic study of knowledge management in primary care. *Bmj*, 329(7473), 1013.
- Gill, B., Coffee-Borden, B., & Hallgren, K. (2014). *A Conceptual Framework for Data-Driven Decision Making*. Mathematica Policy Research.
- Hannaway, J. (1989). *Managers managing: The workings of an administrative system*. Oxford University Press.
- Honig, M. I. (2003). Building policy from practice: District central office administrators’ roles and capacity for implementing collaborative education policy. *Educational Administration Quarterly*, 39(3), 292–338.
- Honig, M. I. (2006). Street-level bureaucracy revisited: Frontline district central-office administrators as boundary spanners in education policy implementation. *Educational Evaluation and Policy Analysis*, 28(4), 357–383.
- Ikemoto, G. S., & Marsh, J. A. (2007). chapter 5 Cutting Through the “Data-Driven” Mantra: Different Conceptions of Data-Driven Decision Making. *Yearbook of the National Society for the Study of Education*, 106(1), 105–131.
- Ingram, D., Seashore Louis, K., & Schroeder, R. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *The Teachers College Record*, 106(6), 1258–1287.
- Jennings, J. L. (2012). The effects of accountability system design on teachers’ use of test score data. *Teachers College Record*, 114(11), 1–23.
- Karen Seashore Louis. (2006). Changing the Culture of Schools: Professional Community, Organizational Learning, and Trust. *Journal of School Leadership*, 16(No. 5), 477–489.
- Kennedy, M. M. (1982). *Working Knowledge and Other Essays*.
- Kerr, K. A., Marsh, J. A., Ikemoto, G. S., Darilek, H., & Barney, H. (2006). Strategies to promote data use for instructional improvement: Actions, outcomes, and lessons from three urban districts. *American Journal of Education*, 112(4), 496–520.
- Khilnani, S. (1993). *Arguing revolution: the intellectual left in postwar France*. Yale University Press.
- King, J. A. (1988). Research on evaluation use and its implications for evaluation research and practice. *Studies in Educational Evaluation*, 14(3), 285–299.

- Lachat, M. A., & Smith, S. (2005). Practices That Support Data Use in Urban High Schools. *Journal of Education for Students Placed at Risk (JESPAR)*, 10(3), 333–349.
- LAUSD Pilot Schools. (n.d.). Retrieved June 23, 2014, from [http://pilotschools.lausd.net/apps/pages/index.jsp?uREC\\_ID=190037&type=d&pREC\\_ID=393572](http://pilotschools.lausd.net/apps/pages/index.jsp?uREC_ID=190037&type=d&pREC_ID=393572)
- Leviton, L. C., & Hughes, E. F. X. (1981). Research On the Utilization of Evaluations: A Review and Synthesis. *Evaluation Review*, 5(4), 525–548.
- Levitt, R. (2003). GM crops and foods. Evidence, policy and practice in the UK: a case study. *ESRC UK Centre for Evidence Based Policy and Practice, Queen Mary University of London, Working Paper*, 20.
- Light, D., Wexler, D., & Heinze, J. (2004). How practitioners interpret and link data to instruction: Research findings on New York City Schools' implementation of the Grow Network. In *annual meeting of the American Educational Research Association, San Diego, CA*.
- Lincoln, Y. S. (1985). *Naturalistic inquiry* (Vol. 75). Sage.
- Lindblom, C. E., & Woodhouse, E. J. (1968). *The policy-making process* (Vol. 4). Prentice-Hall Englewood Cliffs, NJ. Retrieved from <http://library.wur.nl/WebQuery/clc/1850663>
- Liphart, A. (1975). The Comparable-Cases Strategy in Comparative Research. *Comparative Political Studies*, 8(2), 158–177.
- Little, J. W. (2007). chapter 9 Teachers' Accounts of Classroom Experience as a Resource for Professional Learning and Instructional Decision Making. *Yearbook of the National Society for the Study of Education*, 106(1), 217–240.
- Little, J. W., Gearhart, M., Curry, M., & Kafka, J. (2003a). Looking at student work for teacher learning, teacher community, and school reform. *Phi Delta Kappan*, 85(3), 184–192.
- Little, J. W., Gearhart, M., Curry, M., & Kafka, J. (2003b). Looking at student work for teacher learning, teacher community, and school reform. *Phi Delta Kappan*, 85(3), 185.
- Los Angeles Unified School District. (2012). *Pilot Schools Accountability System Guide*.
- Mahoney, J. (2000). Path dependence in historical sociology. *Theory and Society*, 29(4), 507–548.

- Mandinach, E. B., Honey, M., & Light, D. (2006). A theoretical framework for data-driven decision making. In *annual meeting of the American Educational Research Association, San Francisco, CA*.
- Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). Making sense of data-driven decision making in education.
- Martinez, R. A., & Quartz, K. H. (2012). Zoned for Change: A Historical Case Study of the Belmont Zone of Choice. *Teachers College Record*, 114(10).
- Mason, S. (2002). *Turning data into knowledge: Lessons from six Milwaukee public schools*. Madison, WI: Wisconsin Center for Education Research.
- McDougall, D., Saunders, W. M., & Goldenberg, C. (2007). Inside the black box of school reform: Explaining the how and why of change at Getting Results schools. *International Journal of Disability, Development and Education*, 54(1), 51–89.
- Means, B., Padilla, C., DeBarger, A., & Bakia, M. (2009). Implementing Data-Informed Decision Making in Schools: Teacher Access, Supports and Use. *US Department of Education*.
- Means, B., Padilla, C., & Gallagher, L. (2010). Use of Education Data at the Local Level: From Accountability to Instructional Improvement. *US Department of Education*.
- Merriam, S. B. (1998). Qualitative Research and Case Study Applications in Education. Revised and Expanded from “Case Study Research in Education.”. ERIC.
- National Center for Learning Disabilities. (n.d.). What is Response to Intervention (RTI)? | RTI Action Network.
- Nutley, S. M. (2007). Using evidence: how research can inform public services. Bristol, U.K: Policy Press.
- Park, V., & Datnow, A. (2009). Co-constructing distributed leadership: District and school connections in data-driven decision-making. *School Leadership and Management*, 29(5), 477–494.
- Patton, M. Q. (2005). *Qualitative research*. Wiley Online Library.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th Edition). Sage.
- Quartz, K. H., Kawasaki, J., Sotelo, D., & Merino, K. (2014). Supporting assessment autonomy: How one small school articulated the infrastructure needed to own and use student data.

- Journal of Educational Change*, 15(2), 125–152. <http://doi.org/10.1007/s10833-013-9219-4>
- Quinn Patton, M. (1988). Six honest serving men for evaluation. *Studies in Educational Evaluation*, 14(3), 301–330.
- Ragin, C. C. (1994). Introduction to qualitative comparative analysis. In *The Comparative Political Economy of the Welfare State* (pp. 300–309). Cambridge University Press.
- Richards, L. (2009). *Handling qualitative data: A practical guide*. Sage Publications.
- Rickinson, M. (2005). Practitioners' use of research. In *National Education Research Forum Working Paper* (Vol. 7).
- Rock, T. C., & Wilson, C. (2005). Improving teaching through lesson study. *Teacher Education Quarterly*, 77–92.
- Saldaña, J. (2015). *The coding manual for qualitative researchers*. Sage.
- Saunders, W. M., Goldenberg, C. N., & Gallimore, R. (2009). Increasing achievement by focusing grade-level teams on improving classroom learning: A prospective, quasi-experimental study of Title I schools. *American Educational Research Journal*, 46(4), 1006–1033.
- Scholz, R. W., & Tietje, O. (2002). *Embedded Case Study Methods - Integrating Quantitative and Qualitative Knowledge*. Thousand Oaks, CA: Sage Publications, Inc.
- Scott, W. R. (1981). *Rational, natural, and open systems*. Prentice-Hall, Englewood Cliffs.
- Scott, W. R., & Davis, G. F. (2015). *Organizations and organizing: Rational, natural and open systems perspectives*. Routledge.
- Simons, H., Kushner, S., Jones, K., & James, D. (2003). From evidence-based practice to practice-based evidence: the idea of situated generalisation. *Research Papers in Education*, 18(4), 347–364.
- Small, M. L. (2009). 'How many cases do I need?': On science and the logic of case selection in field-based research. *Ethnography*, 10(1), 5–38.
- Spillane, J. P. (2012). Data in Practice: Conceptualizing the Data-Based Decision-Making Phenomena. *American Journal of Education*, 118(2), 113–141.

- Spillane, J. P., & Miele, D. B. (2007). chapter 3 Evidence in Practice: A Framing of the Terrain. *Yearbook of the National Society for the Study of Education*, 106(1), 46–73.
- Stecher, B., Hamilton, L. S., & Gonzalez, G. (2003). *Working smarter to leave no child behind*. Rand Corporation. R
- Los Angeles Unified School District. (2013, May). LAUSD Teaching and Learning Framework. Retrieved from [http://achieve.lausd.net/cms/lib08/CA01000043/Centricity/Domain/88/2013%202014%20TLF%20Booklet\\_FINAL.pdf](http://achieve.lausd.net/cms/lib08/CA01000043/Centricity/Domain/88/2013%202014%20TLF%20Booklet_FINAL.pdf)
- Thorn, C. A. (2001). Knowledge Management for Educational Information Systems. *Education Policy Analysis Archives*, 12, 47.
- Timperley, H. (2009). Evidence-informed conversations making a difference to student achievement. In *Professional learning conversations: Challenges in using evidence for improvement* (pp. 69–79). Springer.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207–232.
- Tversky, A., & Kahneman, D. (1975). Judgment under Uncertainty: Heuristics and Biases. In D. Wendt & C. Vlek (Eds.), *Utility, Probability, and Human Decision Making: Selected Proceedings of an Interdisciplinary Research Conference, Rome, 3–6 September, 1973* (pp. 141–162). Dordrecht: Springer Netherlands.
- Wayman, J. C. (2007). Student data systems for school improvement: The state of the field. In *TCEA educational technology research symposium* (Vol. 1, pp. 156–162). ProActive Publications Lancaster, PA.
- Wayman, J. C., Conoly, K., Gasko, J., & Stringfield, S. (2008). Supporting equity inquiry with student data computer systems. *Data-Driven School Improvement: Linking Data and Learning*, 171–190.
- Wayman, J. C., Stringfield, S., & Yakimowski, M. (2004). Software enabling school improvement through analysis of student data. Report.
- Weiss, C. (1995). The four“ I’s” of school reform: How interests, ideology, information, and institution affect teachers and principals. *Harvard Educational Review*, 65(4), 571–593.
- Weiss, C. H. (1982). Policy research in the context of diffuse decision making. *The Journal of Higher Education*, 619–639.

- Weiss, C. H. (1988). Evaluation for decisions: Is anybody there? Does anybody care? *Evaluation Practice*, 9(1), 5–19.
- Weiss, C. H. (1998). Have We Learned Anything New About the Use of Evaluation? *American Journal of Evaluation*, 19(1), 21–33.
- Weiss, C. H. (1999). The interface between evaluation and public policy. *Evaluation*, 5(4), 468–486.
- Wood, M., Ferlie, E., & Fitzgerald, L. (1998). Achieving clinical behaviour change: a case of becoming indeterminate. *Social Science & Medicine*, 47(11), 1729–1738.
- Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Yin, R. K. (2012). *Applications of case study research* (3rd ed). Thousand Oaks, Calif: SAGE.
- Young, V. M., & Kim, D. H. (2010). Using Assessments for Instructional Improvement: A Literature Review. *Education Policy Analysis Archives*, 18, 19.
- Zeuli, J. S. (1994). How do teachers understand research when they read it? *Teaching and Teacher Education*, 10(1), 39–55.