EVIDENCE OF SYSTEM:

A NETWORK MODEL CASE-STUDY

OF

SEVENTH GRADE SCIENCE ASSESSMENT PRACTICES

FROM CLASSROOMS TO THE STATE TEST

by

Philip John Piety

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Education)
in The University of Michigan
2008

Doctoral Committee:

Professor Joseph S. Krajcik, Co-Chair
Professor Pamela Ann Moss, Co-Chair
Professor Jay Lemke
Professor John Malcolm Swales

To my loving family: Sarah, Dustin, and TQ

Acknowledgements

This research would not have been possible or would have been in a much weaker form had it not been for a number of people whose support and inspiration were essential to my work.

The first person I must thank is Sarah Ball who believed in the possibility of pursuing a doctoral education in the way that I did and in the area I wanted to study. Sarah is the pure scholar in our family and my inspiration. I had not the wisdom to pray that she would come into my life when she did. I know now that someone must have.

The next acknowledgement goes to my committee. Early in my doctoral studies a wise professor advised me to find that one person who I want to emulate in my work. I never did find that one person. But, in my committee I did find the four best scholars who each serve as a role model, representing standards I strive to embody. In alphabetic order, I will forever be indebted to Joseph Krajcik. He is a man of decency who builds communities of wonderful scholarly exchange. I have been privileged to meet and learn from many scholars in the forums he has initiated and from him in many teachable moments that stealthily pass. If I were to have the honor and responsibility of students of my own at some point, I would emulate him and his approach. Professor Jay Lemke is an intellectual traveler and has been for me a tour guide. From oracle bones to thinking space, he has shown me the importance of quality scholarship and introduced me to different communities outside of education where thought and theories, often beyond my ken, help make the world a clearer place. His influence is in every chapter of this work. Professor Pamela Moss has shown me the best example leadership for educational research. Her work to build bridges across communities and discourses are making our field stronger from its foundations and her work in this are makes me proud to be an educational researcher. She is a passionate advocate for dialogue and appreciating the other's perspectives. I will forever be grateful she joined my effort and I continue to hear her voice in the research decisions I make. Professor John M. Swales is someone I know best, appropriately, through texts. He creates living worlds in simple bound stacks of paper where texts that real people write, read, fashion, and hold are alive. His thaumaturgy in explorations of symbolic and social practice is the goal that this work strives to emulate.

I have had the benefit of learning from a wonderful faculty at the University of Michigan who shaped my thinking about education and research. My deepest gratitude goes to Annemarie Palincsar whose care for all students, especially me, was evident in all she did. I think of her as an important guide in what led to this research. Magdalene Lampert was also an important influence on me and my thinking about theory and practice.

During the course of my doctoral studies, I was fortunate to have a number of wonderful opportunities to learn about science education and its relationship to policy and practice. Colleagues at the Center for Highly Interactive Classrooms, Curricula, and Computing (hi-ce^3) were very helpful for me, especially the friendship of Jay Fogleman who could teach me more about educational practice in a short conversation than many graduate courses and Cesar Delgado, who shared many important insights about educational research and family life. As a special member of the Center for Curricular Materials in Science, I benefitted tremendously from the conferences and as a witness to the collaborations that occurred around changing science educational practice. I also benefitted from a summer with the American Association for the Advancement of Science’s Project 2061 and thank George DeBoer and Jo Ellen Roseman and their team for making that such an important learning experience for me. I also am indebted to Andy Anderson and colleagues at Michigan State for teaching me about science education and learning progressions in real-world practice.

The writing of this dissertation was a difficult process. Charles Piety was very helpful in lending his knowledge of science and exposition to some of the more difficult sections as they were coming out of the very rough drafting stage. Providentially, Linda O’Brien assisted in the last several weeks of this effort and helped me to conceive of how this work fits within contemporary discourse. Finally, I thank Wilson and Molly. They know the value of a purr and a pet at the most important of times.

Table of Contents

List of Figures

List of Tables

List of Documents

List of Photos

List of Transcripts

List of Appendices

Abstract

With science education in the United States entering a period of greater accountability, this study investigated how student learning in science was assessed by educators within one state, asking what systemic assessment approaches existed and how the information from them was used. Conducted during the 2006-2007 school year, this research developed and piloted a *network-model case study* design that included teachers, principals, administrators, and the state test development process, as well as several state-level professional associations. The data analyzed included observations, interviews, surveys, and both public and private documents. Some data were secondary. This design produced an empirical depiction of practice with a web of related cases. The network model expands on the hierarchical (nested) models often assumed in the growing literature on how information is used in educational contexts by showing multiple ways in which individuals are related through organizational structures.

Seven case study teachers, each employing assessment methods largely unique and invisible to others in their schools, illustrate one set of assessment practices. The only alternative to classroom assessments that could be documented was the annual state accountability test. These two assessment species were neither tightly coupled nor distinct. Some teachers were partners in developing state test instruments, and in some cases the annual test could be seen as a school management resource. *Boundary practices* -- activities where these two systems connected -- were opportunities to identify challenges to policy implementation in science education. The challenges include standards, cognition, vocabulary, and classroom equipment. The boundary practices, along with the web of connections, provide the outlines of potential (and often unrealized) synergistic relationships.

This model shows diverse indigenous practices and adaptations by actors responding to pressures of change and persistent historical tensions of diversity and control. It provided evidence of a broadening instructional agenda and rapid deployment

of information infrastructures for collection, dissemination, and analysis of student information. The model became a lens to view these changes and paths that policy for science education may take for implementation. It also became a lens to evaluate accountability policies to see how models embedded within policies may fit with current practice.

Chapter 1 Introduction

In response to both increasing public accountability for science education in the United States and growing interest in the ways that educational organizations perform as systems for learning, this study sought to understand how information from assessments of student learning in science was used within a particular set of educational contexts. It is a study that analyzes systemic characteristics from a number of perspectives, looking at collections of related individuals and organizations, at practices and artifacts, and at reciprocal relations.

Science education presented an important research problem and opportunity for this study. At the time it was conducted, the federal No Child Left Behind (NCLB) legislation had been active for several years, requiring states to measure students' performance on two subjects: math and literacy. Schools were held accountable for changes in student scores from year-to-year[1] for these two subjects in what NCLB defined as Adequate Yearly Progress, or AYP (NCLB, 2002). Initially, states were not required to test for science, but NCLB required all states to begin implementing science testing at the start of the 2007-2008 school year, and there was consideration in policy circles that science education might be elevated to an AYP subject in the reauthorization of the controversial law. Furthermore, science education and how it is assessed had gained attention in national research and policy circles with the release of a series of studies conducted by the National Research Council fostering expectations that more systematic approaches to science education could be near. Despite this interest in science education as a more publicly accountable (systemic) activity, a paucity of research existed that documented current practice in this area of emerging importance.

[1] While AYP may be thought of as a measure of student learning, it is actually applied to measurements of different groups of students across years.

While conceiving of educational practice in systemic terms is a common feature in policy and research, the actual systemic character of educational organizations and group activity is an area of active questioning (Cohen, 1995). How collections of individuals and organizations behave in coherent, reciprocal ways in actual practice in education has not often been demonstrated empirically. While primarily an investigation into science education and assessments, this study is also a description of indigenous practice from a systemic perspective. It is bounded within a single U.S. state -- a geographic area that NCLB and other federal policies designate should be considered a system. It considers practitioners from local classrooms to individuals and state-level organizations, while focusing on the seventh grade. It therefore acts as a topical investigation delineated by policy and geography in order to develop an understanding of the systematic character of a portion of the U.S. educational enterprise.

While the notion of a system is easily deployed in theoretical and policy discussions, it exposes significant and persistent questions about how educational processes can be managed, predicted, and controlled. Before introducing research questions, it is important to acknowledge the diversity of applications for the term "system." Where organizations are defined in political and geographic terms, the system is defined conceptually. While a system is generally considered to be a collection of interrelated and often independent parts that comprise a coherent whole, as in a railroad system or a tissue system, the focus of the term can vary. A system can be described in terms of objects and artifacts, practices and processes, and behaviors and actions where systemic linkage can be observed, or some combination of these. Essentially, it is interconnections in action that constitute systematicity, and terms that describe these connections -- "linkage," "relationship" and "coupling" -- all convey a sense of this reciprocal nature.

The American visionary R. Buckminster Fuller is reported to have said that "synergy means behavior of whole systems unpredicted by the behavior of their parts taken separately." This study is organized by a related concept: systematicity, the notion that different parts of some larger whole operate together in ways not evident from their individual components' operation. The notion of an educational system had been an important feature of the policy climate existing for several decades prior to this work.

Within the current era of comprehensive school reform, educational systems and their constituent schools, leaders, teachers, and students have been subjected to increased measurement and accountability according to explicit standards. The conception of an educational system encoded in policy, most prominently NCLB, has driven this investigation. Since information about student performance flowing within and across educational organizations is an essential systemic component of NCLB, that flow of information became the guiding analytic object in this investigation.

Considering educational practice as a system also entails the notion that it can operate in an asynchronous, non-synergistic, or loosely coupled manner (Weick, 1976). An internal combustion engine could be considered a system. When it behaves properly, the system starts when directed and generates mechanical forces in response to various controls. When it is not properly functioning, it may fail to start, may respond incorrectly by hesitating when the gas pedal is pressed, or may run for a period before overheating. In contrast, systemic failure in educational contexts can occur even as the system operates smoothly. Systemic failure of an educational system can be seen as a type of synergistic failure where the system runs, but the intended outcomes are not reached. Synergistic failure in education might mean that classroom instruction continues, but without a strong relation to other components, such as leadership direction or performance standards.

Researchers have posited that information from assessments of student learning could be used to inform a larger educational system (Frederickson & Collins, 1989; Pellegrino, Chudowsky, & Glaser, 2001). I am, then, not treating assessments as a technical matter between the student and the test or viewing them from an accountability perspective to evaluate educational effectiveness. Rather, I am considering them as informational resources with the potential to be used across contexts and boundaries of responsibility (Daft & Lengel, 1986). Conversely, they also have the potential not to be used and to accumulate within organizations without significant use (Feldman & March, 1981).

In considering assessments as informational resources, I focus on the artifacts and how they are used organizationally. Assessment artifacts include the tasks given to students and the derivative information from those tasks, such as score reports and grades. Artifacts can also be considered to include activity structures of use that animate these tasks and texts (Halverson, 2003). I include student performance information that comes from a broad range of tasks, seeking to understand how that information is used, by which professionals and for what purposes related to information. One important set of task structures I consider is the annual assessment program from the state. The study of this annual accountability system includes evidence from both its site of production, as managed by the state government and its contractors with participation from local practitioners, and also at sites of its consumption in local schools.

Research Goal: A Systemic Depiction as a Scientific Model

In this study, I operationalize the "system" as a way of framing research that occurs across related organizational entities. With this move, the study becomes in some respects a kind of systemic investigation -- an investigation with the potential to offer some contribution to the understanding of educational systematicity from one small aspect of educational practice operating over a particular organizational landscape. This emergent study began with two basic questions that were framed by the policy climate of NCLB, where data and assessments can be seen not only as diagnostic tools, but also as relevant resources for decisions that educators make in supporting students:

1. **What coordinated systems (practices and artifacts) for assessing student learning were used for seventh-grade science education?**

2. **Who used these systems, and what boundaries did these systems cross relevant to instruction?**

These two questions initially focus the investigation on student learning and instruction. Excluded from the investigation are systems that assess other aspects of student life, for example demographics, commitment to school, and interests. Also excluded from this study are the uses of student performance information in ways not related to student learning or instruction, such as for school accountability, school choice, and funding decisions. While all of these areas are important tenets of NCLB, the focus of this research is on the instructionally relevant uses of assessment information. Each of these two questions exposes different ways that the term "system" can be used in this targeted context. The first question considers systems as artifact (text) centered entities with practice dimensions. The second question considers the individuals and organizations connected by artifacts. These two questions elicit a third question that considers the organizational nature of the area the first two questions are focused upon:

3. **What is the organizational environment, in terms of structures and boundaries, under which these systems were operating?**

By focusing on one subject, in one grade, and across organization types (and in one state), I am able to ground this study in an understanding of what occurs within the core pedagogical location of classrooms and then proceed through levels of aggregation, including to the state's development of the annual accountability tests. This approach can be thought of as a vertical sampling strategy, although the depiction from this study shows the system to be other than hierarchical, to have multiple paths in a network structure. While the guiding focus is on instructionally-relevant assessments, this research ventures beyond the classroom and seeks to understand assessment issues that exist over an organizational domain. This domain is where policy begins to be translated into practice. The organizational domain is also where many traditional entities (teachers, school leaders, curriculum specialists, and administrative structures) are currently subjected to redefinition pressures related to standards and accountability. Traditional boundaries of responsibility are becoming more fluid, with the linking of leadership and teaching in a signature realignment that is relevant to this study (Copeland & Knapp, 2006). No longer are teachers and administrators easily placed into instructional and management roles. Now both have instructional responsibility and both can be seen as leaders in the community of a school (Spillane, 2006).

The answers to these three questions as derived from this study can be expressed in a number of ways, including in systemic terms. When taken as discrete answers to the questions, the results can form a series of propositions. When assembled into a whole, the results can operate as a systemic depiction, an empirically-developed model that can be used not only in propositional terms, but also in a policy dialogue. The term "depiction" is used deliberately, to contrast with terms like "representation," which has general uses and uses in research discourses that indicate a kind of general similarity between the model and its referent. “Depiction” here is used in the way that Tversky (1981) used it: to refer to cognitive images of space that she showed the human mind maintains which are distorted from the reality they refer to. Much in the way that a subway map will emphasize certain features and diverge from spatial fidelity for ease of use, mental depictions support individual cognitive processes through a highlighting of certain features rather than by seeking to present an exact image of an external whole.

While the relationship between symbols or symbolic assemblages and underlying meaning is complex, contingent, and beyond the scope of this research, the terms that are used matter in framing the research. The term “depiction” is used in this context to indicate a descriptive approach that abstracts and highlights a more complex reality in purposeful ways, rather than attempting to replicate it. This study aims to produce a depiction that highlights features relevant to assessment in educational practice. This depiction, then, is not intended to achieve universal similitude or to be used as the ideal meta-representation of educational practice. Rather, it is a biased view, structured by the policy climate, including NCLB, which places primary emphasis on student assessment information.

Another important semiotic relationship between what this study aims to produce and the research process is the conception of the depiction as a *model*. Natural scientists use models as central mediational artifacts. Ecologists will document relationships between organisms which produce food and those which consume food in an ecosystem; paleontologists might represent the evolution of features in a species over various epochs; climatologists express various significant features of a weather front using symbols and notations. While each of these disciplines has certain representational forms, they are all used to make visible an underlying model of reality that is not readily visible. In the same way, the products of this study can be considered as elements of a model that make visible certain underlying systemic characteristics that would not be apparent to the naked eye. The model this study produces exhibits properties that span time and location, showing underlying principles of behavior in response to policy, specifically NCLB and its construction of assessment information. This model then becomes a key systemic construct both for evaluating school outcomes and supporting decisions made by school leaders.

Policy Climate and Relevant Literature

As a systemic investigation, this research gains significance from several contemporary trends in policy and research. Federal mandates in NCLB required states to begin testing students in science in the 2007-2008 school year (NCLB, 2002). NCLB applies systemic pressure on schools based on student test performance. Contemporary discussions related to its reauthorization indicate the possibility that science education will have an even greater role in determining a school's ability to meet performance standards. Furthermore, recent national forums promoting a new set of approaches to large-scale assessment of science education, including the National Educational Assessment Program, or NAEP (National Assessment Governing Board, 2006) and the related National Research Council publication *Systems for State Science Assessment* (Wilson and Bertenthal, 2005), are organizing around cross-grade curricular definitions called *learning progressions.* Through grade-spanning assessment targets (constructs), learning progressions may exert pressures for vertical (cross-grade) organizational coupling. In addition, recent research that I review in this dissertation indicates that information about student performance from various assessments was being used in a number of consequential practices in educational organizations involving different practitioners, including leaders, administrators, teachers, and specialists. The cross-role or cross-organization interactions represented in this new literature exemplify what I consider to be examples of systemic activity related to information use.

NCLB advanced two important conceptions of student assessment information. First, student assessments became the primary outcome variable that schools became measured on. Schools were required to report student performance in year-over-year results by subgroup so that rather than being held accountable for total performance they were held accountable for performance of different groups provided the number of students in that group reached a minimum threshold. Second, NCLB, emphasized districts base their decisions on evidence and data. This part of NCLB seems to be often interpreted similar to school outcome to mean student performance data although it is possible that other forms of information including information from research could be used as well.

While there are expectations and visions for a systemic science education, one that uses information across grades and where different professionals can play active roles in ensuring broad student success, the research base documenting actual practice in science education is thin. Some, including Love (2002) and Boudett, City, and Murnane (2005), present a vision of practice where assessment information can guide instructional decisions, as indicated in the technical language of NCLB itself. However, my review of the relevant literature failed to show a solid foundation for these expectations. Few of the publications in this emerging literature attend specifically to science education. Most are centered on math and literacy (the two NCLB accountability subjects) and/or studies drawn from specific programmatic reform efforts, rather than indigenous practice.

The emerging literature on how information/data are used to support collaboration and systemic activity in education shows a variety of research approaches. While this literature provides important parts of a foundation for this study, it also lacks many of the norms found in mature research genres and raises many important questions about the best ways to study these processes. Most of this research is limited to a small part of the organizational spectrum that I name *systemic aperture.* Also, many of the studies I review present only cursory discussions of their research methods. The data from these studies are often drawn from another study or from very brief observations. I conclude that the literature cannot yet be considered what Kuhn (1970) would call a *normal science*.

While having many limitations in terms of its ability to frame the study, the literature indicated new ways of thinking about educational practice. The emerging literature draws on many different types of intellectual influences, including policy, administration/organization studies, teaching and learning, and also on other recent studies into information use in education showing a developing professional community. This literature indicates that some more advanced practices for information use may involve using multiple data sources, longitudinal views of students over time, and intermediate or interim assessments that track progress during a school year.

This study responds to these gaps and opportunities in the literature by providing a systemic account of current practices that indicates the types of situations that may be encountered in routine activities and by common practitioners.

Context & Method

Systemic character cannot be measured in a universally formulaic manner. If we begin with Fuller's position that synergy in a system is unpredicted by the behavior of its parts, then the study of small sets of individual elements may be limited by the elements' capacity to represent systemic character. When systemic components are drawn only from one localized sector of the theorized system (ex: schools within districts), they can be limited in their ability to show dialectical and reciprocal systemic relations. Conversely, when they are selected from a broader set of organizational elements, then it may be more possible to see systemic relations where those elements are related theoretically and/or empirically.

The approach I took in this study, both in scope and method, was influenced by policy as well as by theoretical and practical considerations. The current regulatory landscape of education in the United States makes the individual state an important unit of analysis. NCLB delegates to the individual states many important implementation details, and states are responsible for the development and administration of performance standards and tests required by NCLB. Accordingly, I limited my study to one of the fifty states. Within this state I studied several levels of educational organizations, from the state government to five individual schools and seven classrooms. Between schools and the state were a variety of intervening structures and organizations that I also included in this study. While crossing these different contexts, I strove for unity in both the topic (seventh grade science) and research method. The state I studied, Michigan, has a decentralized organizational structure with many small districts and a layer of intermediate school districts, or ISDs, that provide a range of support services to the local districts. There is also a network of Math/Science Centers that provide support for science education. These centers are related to ISDs, but not in a consistent way. Across the state during this time, there were both funding shortages for education related to the state's economic condition and new sets of science standards being developed, but not yet in use.

There were practical questions that a systemic study raises, as well. Can a researcher go to a location and observe systemic activity in contiguous units? In organizational action, in contrast to a classroom, the preponderance of the communication occurs asynchronously, across time and context, and through documents or texts rather than in direct encounters where the textual ecology is shared and local. While "the system" may be continuously engaged, it is working on many different topics, often in temporally separated contexts. Activities where assessment information is used could occur at different times and in different locations. How can research that is interested in asynchronous hetero-locality practices, and yet cannot be everywhere at all times, efficiently and accurately document practices in sufficient detail to understand organizational action?

The methodology developed for this study is an adaptation of the case study method (Yin, 2002) drawing off of Actor-Network Theory or ANT (Latour, 2005; Law, 1992). ANT comes from an ethnological tradition and calls for close study of particular contexts. Unlike many ethnologies or ethnographies that focus on a single locus in a place or a culture, ANT studies across contexts. ANT applies similar methods to the macro and the micro and considers artifacts and technologies to have potential to influence events in the same way that human actors do. Important considerations in this research are issues of temporality and understanding of activities that exist in multileveled configurations with some shared historic context (Lemke, 2000). Swales' (1998) textography method, which uses practice documents as indicators of patterns and temporal relationships, contributed to prioritizing texts of practice as important units of analysis and triangulation. Throughout the analyses, texts are used to look for confirmatory/dis-confirmatory evidence of events not observed. This allows the study to retain a consistent method while traversing different contexts or levels of the educational system. It is an inductive approach, as the sampling strategy of focusing on one class of information, for one curricular topic and one grade level provided an avenue for study that revealed the system from the inside out and provided important details about the organizational structure particular to this research context.

Preview of Five Analytic Steps

Not only does this study contain many different kinds of data, often localized to specific cases; it also uses many tactical and interrelated analyses. This data set presents many options for displaying findings because there is no predetermined hierarchy in the network structure (Latour, 2005). The approach I selected features five analytic steps. Each of these analyses presents a distinct portrait, a lens onto organizational/systemic dimensions of educational practice. I begin with the aspects closest to the day-to-day lives of students and end with the issue of change across the network.

The first analysis focuses on classroom practice. In taking a systemic view of what teachers do with assessments, I found each teacher in the study had a coordinated assessment approach where different types of instruments, such as quizzes and tests, played specific roles. From teacher to teacher these roles were often similar, but did not present a uniform or common systemic pattern. Rather, each teacher's practices formed a unique system/subsystem. These systems were developed through individualized approaches where teachers are *bricoleurs:* they draw on available resources in textbooks, the Internet, the work of colleagues, and popular media. As with instruction methods, the assessment of seventh graders is sensitive to their human development and maturity. If we consider classroom assessment as a communication system embedded within a type of social conversation between the teacher and the class (Yinger, 1990), the discourse has semiotic tasks that have multiple, rather than singular purposes (Scollon, 2001). Some of these purposes include building knowledge, teacher self-evaluation, and providing multiple paths for success, in addition to measurement. The students of the teachers in this study often had opportunities to do projects where expressions of identity and invention of representations, also known as meta-representational competence (DiSessa, 2005), were encouraged. The emphasis, however, that different teachers placed on different types of tasks and skills varied significantly.

Some teachers in this study were also beginning to use advanced assessment technologies in ways compatible with large-scale assessment information architectures, although they did not seem to be doing so as part of an administrative mandate, but rather organically to meet classroom needs. However, not all teachers in this study had the same flexibility or opportunities. Some teachers worked in schools that used a science kit approach. These teachers were presented with the resources for instruction and a script or guide in the manner commonly used for elementary teachers, whose knowledge of, and even familiarity with science tends to be much lower than that of science teachers in higher grades.

Each teacher in this study had an assessment practice that could be characterized by a personalized temporal signature with intervals or timescales (Lemke, 2000) associated with specific assessment instruments. The portrait of assessment practices inside seventh-grade science classrooms that emerged for most teachers is one where students encounter a continuous stream of tasks with assessment potential as the teacher progresses through the curriculum. This continuous, short-timescale model is compatible with the expectations of many in the educational system, including school leaders and testing professionals. For most teachers in this study, the temporal interval or *timescale* (Lemke, 2000) was more prominent than the role of one interval over another, the *timeframe*. The kit teachers, however, had a temporal structure imposed upon them by the distribution schedule for the kits. Since kits are shared within a region, the kit teachers' work was structured by when the kits arrived and when they needed to be sent on. This physical kit resource compromised the district's capacity for topical synchronization across classrooms as might have been helpful for certain common intermediate assessments, because the same pool of kits would rotate through different classrooms in sequence.

A second analysis focused on one professional association and two individuals who were influential in it. The association is the leading science teachers' professional community in Michigan and one that has many connections to other study data. This organization supported a survey of its membership that I designed to understand their practices of assessment and their experiences in a range of areas, including the annual MEAP test and, for some, their experience in helping to develop the test. This analysis looks not at typicality in terms of science educators, but rather at uniqueness. In comparing the science teachers in the study to the survey respondents, I find the case study teachers are not unique in their practice, although I do not go so far as to find any of them typical. Then in looking at the science teacher association leaders, who are cases in the study, I leverage their unique positions and experiences to develop deeper understandings of historical progression and intermediate layers. These high-profile, highly networked leaders were able to provide insights that less unique science educators (ex: the case study teachers) were not able to provide.

In a third analysis, the focus shifts further, from science education to the study's full network of participants, which included 50 individuals and 19 different organizations. In this analysis, I highlight the interconnections between these individuals and organizations. In several cases I show how some individuals play roles that Wenger calls brokers (Wenger, 2000) and others call boundary spanners (Honig, 2006; Tushman & Scanlan, 1981). I also explore the artifactual connections between the cases using the perspective of boundary objects (Star and Griesemer, 1989). The case structure allows a view of practice networks where the individual roles of actors are heterogeneous, rather than equivalent. Within this study, some teachers have roles that are more isolated from the network, while others are more connected and have a greater visibility and network influence. Often in educational network analysis, collections of individuals with common job titles (ex: teachers) are homogenized, masking what this study shows is an important factor in understanding patterns of influence: organizational affiliation. This analysis also looked across the network for positions that individuals held regarding accountability and found that this study's data did not support the same type of leveled epistemology reported in some recent research. This analysis also confirms the predominance of two assessment systems across the state for middle school science: classroom tasks and the state test.

The fourth analysis looks into the network at two types of boundary practices (Wenger, 2000) where the state and classroom systems can be seen to interact. This analysis draws chiefly on five review sessions, three in the state and two in schools, where the practitioners discuss the state test items under development and local test results, respectively. This analysis confirms the role of the individual assessment item as a central mediator. The individual item, which Swales (1998) refers to as a *microtext*, takes on a role seemingly more important than alternative candidates of either the learning standard or the item group called a strand (ex: earth science, life science, reflecting on scientific knowledge), causing the individual item to appear as the principal boundary object (Star & Griesemer, 1989) helping to organize group activity. This analysis also highlights asymmetries in the dialogical nature of these activity systems. The point where the classroom and accountability perspectives intersect is designed and managed by the state. Those practitioners who participate report many benefits in terms of their growth and career development. Even though the state is largely managing the process, it is beneficial for the individual teachers who participated. This analysis also highlights opportunities for alignment between these two systems, such as identifying task discrepancies and addressing the significant role of classroom equipment in students' opportunities to learn about important aspects of scientific measurement.

This fourth analysis also emphasizes the importance of timeframe, as it shows how the school's periodic topology is influenced by what occurs in the higher-level state test cycle. The tests are given early in the year and then the results are released later, after enough of the tests (including alternative and makeup tests) are administered to calculate state comparison and AYP figures. While the interpretive processes differed, each school's processes were largely temporally aligned to the release of the test results. The school year in the individual schools can be divided into segments based on the state test process interaction points, although a subset of survey respondents reported that the test had no impact on their work.

The fifth and final analysis looks at change. It considers the network from a diachronic perspective. Here, the professional association case studies carry a greater load in terms of evidence, as their public documents extend back several years and are more broadly representative of what is occurring within the state. The state testing operation was in a period of transition. It was both developing new science standards and involved in efforts to develop architecture standards for the districts and intermediate assessment tools for schools, beginning with high schools. A combination of evidence from the individual school cases, the nascent literature on information use in education reviewed in this dissertation's second chapter, and proceedings from the professional associations on assessments made it possible to tentatively sketch the outlines of a developmental process for information literacy. That process could explain many differences in the boundary practices reported by the case study schools where the local interpretation of the MEAP results was observed. The theme of topical alignment is clear across the state, because in both individual schools and professional associations, the current period is marked by teachers needing to teach subjects that are aligned with the state test. The significant efforts to develop the kind of infrastructure for information management being undertaken across the state – which much of the available literature indicates is essential for successful use of information supporting educational decisions – is evidence of this important historical shift.

This fifth analysis allows the temporal modeling of the educational enterprise to be extended further. In addition to timescales and timeframes, the analysis exposes episodic periods that do not seem recurrent in the way that segments of school calendars can be shown to be in response to the state test. This analysis also raises important issues related to the diffusion of innovations (Bass, 1969; Rogers, 1991), where there is a process that includes early and later adoption of assessment-related information systems. In educational organizations in this state, however, unlike the situation of consumers of information systems in marketing studies, the adoption of innovations related to student information systems occurs within and is influenced by the state's particular multileveled organizational ecology. The diffusion that was occurring across an irregular organizational landscape brought into focus both important aspects of Rogers' model and also how that model can be complicated by sectors like the decentralized public education system found in Michigan.

Organization of the Document

The document that follows begins with a literature review. Following this is a discussion of some theoretical topics that are useful in understanding the design and analysis of the study. I then present the design, including a formal definition of the methodology, before five analysis chapters. Each of these analyses builds upon the previous ones and follows the order of the five steps discussed above. The conclusion then summarizes key findings and reflects on the placement of this study historically, the limitations and strengths of the evidence, and the contributions the study makes. The short final discussion in the conclusion chapter addresses some ways to extend this line of research.

Chapter 2 Literature Review

The obscurest epoch is today. ~ Robert Louis Stevenson

This chapter sets a scholarly stage for the body of this dissertation. While throughout this document various theoretical, methodological, and empirical resources are used to explicate decisions I have taken, this literature review serves to frame the intellectual community to which the dissertation's products will be primarily addressed (Boote & Biele, 2006). This dissertation joins a relatively new area of educational research, an area with a dynamic and possibly contingent relationship to different professional communities that include those studying: administration, instruction, and technology. Because of the emergent and multidisciplinary character of this area of research, I begin with a discussion on the nature of literature reviews in doctoral dissertations recently featured in *Educational Researcher*. This is a discussion to which I also offer a contribution as I begin my own process of selecting, synthesizing, and constructing the proscenium (Lather, 1999) for the research design and the analyses that follow.

The literature covered in this review is treated somewhat differently than literature reviews in other empirical areas. While the literature – in the form of journal articles, book sections, books, and reports – is reviewed for the questions it can help ask, it is also treated critically and as a data source. The condition and variety of the body of works discussed here support assertions about what may have been broadly occurring in education in the few years prior to the time of this review and the ways that educational researchers are responding to these events. These works help to situate this study historically, as well as suggest the questions that this study can address.

On the Nature of an Educational Research Literature Review

Literature reviews are connected to disciplinary fields. In empirical work, including dissertations, they often situate the questions that the study addresses ino a community context. However, there is no specific formula for what they should include. While an empirical dissertation is a research event, it can also be more, setting the stage for a career that should be a unique contribution to an intellectual field. In discussing the dissertation literature review, Boote and Biele (2005) have argued that it should meet the responsibility of providing *comprehensive coverage* of a field, stating:

> Acquiring the skills and knowledge required to be education scholars should be the focal, integrative activity of predissertation doctoral education. Preparing students to analyze and synthesize research in a field of specialization is crucial to understanding educational ideas. Such preparation is prerequisite to choosing a productive dissertation topic and appropriating fruitful methods of data collection and analysis. (ibid, p.3)

They continue with a discussion grounded in their recent study of dissertations from schools of varying national ranking that, with few exceptions, showed weak treatment of the relevant literature. They say that, “If their dissertation literature reviews are any indication, many of these now-doctors know bits and pieces of a disorganized topic.” (Ibid, p.3.) Part of their solution is a proposal for quality criteria in five categories for this critical dissertation activity. The categories include: coverage, synthesis, methodology, significance, and rhetoric. They then define twelve specific criteria across these categories. In responding to a critique and alternative perspective by Maxwell (2006), they add that the literature review serves a *socializing function*, providing the new scholars with the intellectual community where their work can be situated. The citations in effect form an audience for the work, rather than being purely analytic elements.

Maxwell’s (2006) alternative perspective is that the literature review in a doctoral dissertation is different from a review done by experienced researchers. He views the dissertation as a *research event* where the literature review is crafted as one of several supporting elements in an empirical activity. He advises doctoral researchers to be *selective* in this process, rather than striving for exhaustive coverage. Responding to

Boote and Beile's (2005) description of the literature review as a foundational activity, Maxwell asserts:

> A literature review is an essential tool, and any researcher must learn to use it competently and appropriately, but it is no more the foundation of research than a hammer, or even an entire toolbox, is the foundation of carpentry. (p. 30).

While both perspectives relate to this current work, both are missing important elements for what I will present in this dissertation. Although they provide useful alternative conceptions of the role that the review can play, both underemphasize the evolutionary and multidisciplinary aspects of research communities. While Boote and Beile do include historical contextualization of the field as a quality criterion, they do not discuss the process by which fields exist over time: how they come into existence and, through their life-cycles, go through periods of transformation. Neither speculates about the effects these transformations can have on the type of literature that is available for possible review or how at different points in time, different types of decisions about how to approach the review may be called for.

Mindful that any macro social construction such as a field, society, or actor is inherently imprecise (Latour, 2005), I return to Kuhn's discussion of scientific paradigms (Kuhn, 1970) to discuss what I believe to be two additional dimensions of dissertation literature reviews. The first considers the evolution of a disciplinary field -- its life cycle. Kuhn described the successive nature of scientific regimes as paradigms that succeed each other as sociological processes he termed "scientific revolutions" bring new paradigms into existence, challenging existing scientific communities wedded to older paradigms. Once it is well established, the work within a paradigm is often of a problem-solving nature, which Kuhn referred to as a normal science. Boote and Beile (2005) assert this to be a rare condition in subfields of education and argue that educational researchers must often create the foundation for their work themselves; and the literature review, in their view, is where much of this foundational work is done. Their position places less emphasis on new researchers using frameworks that are already established in a community than Maxwell, whose position may to some extent be more appropriate when a paradigm is established, the science normal, and literatures and genres established.

The second dimension of this dissertation's literature review is the cross-disciplinary connections the literature may entail. It may be that these disciplinary relationships are different when a field is young, when there is a greater need for the field to borrow from other disciplines and attempt to synthesize perspectives that may or may not survive the test of time. Should the ideal of dissertation literature review also include a critical perspective on the field's disciplinary constitution? Should it also ask questions about what other disciplinary perspectives may be beneficial to the field at a particular point in time?

As historians of educational research and practice remind us, educational research has gone through many eras, each having different conceptions of what learning and the educational process are, different notions of teaching and learning, different norms for evidence, and different conceptions of research processes (Gamson, 2007; Lagemann, 2000). Looking back on the last thirty years of educational research, movements such as the "Process-Product" approach from the 1970s, or the or Constructivist theories of learning, and later Social Constructivism came into existence and in some cases left the scene. Fields such as Evaluation. that began in the 1980s and is still active, or the Learning Sciences, that began a decade later, are examples of fields that were aided by important contributions from new technology. In the case of Evaluation, the technology that allows processing of large amounts of data was instrumental in it being possible to perform certain types of analyses efficiently; while Learning Sciences has utilized technology in a range of classroom and learning contexts to support student growth (Edelson, Gordin, and Pea, 1999; Krajcik & Blumenfeld, 2006).

By considering the disciplinary life-cycle perspective – to conceptualize the birth, development, maturity, and transitions within professional communities and how these stages may impact the nature of what an emerging scholar needs to consider – some important synthetic space between may open up. Could it be that the notion of *coverage* that Boote and Biele emphasize would change over time with the accumulation of literature around a given problem? How is the aspect of *selectivity* that Maxwell argues for different when a field is young and there may be fewer coherent positions than later, when more established positions exist?

Even within the life of a particular field, there can be significant shifts in what is considered an evidence base. In *Memory Practices in the Sciences*, Bowker (2006) discusses shifts in the evidentiary processes that have occurred in natural sciences, such as geology, where at one time scientific work was centered on learning how to read the layers of the earth as an historical text, while today much of the scientific work in geology involves the development of databases of evidence that can be shared across researcher settings.[2]. During these evidentiary evolutions, the field of geology itself is consistent. Looking at education with this lens, the attention to using video records in teacher education and the development of video collections (Brophy, 2004; Hiebert, Gallimore, & Stigler, 2002; Pea, Lindgren, & Rosen, 2006) can be seen as similar developments in evidentiary practices enabled by technology. And just as the field of geology has remained as an identifiable professional community through these changes, so has teacher education remained a coherent professional field.

To Boote and Biele's recommendations, I propose that the dissertation should also frame new researchers' work within a depiction of the developmental trajectory of their field. Further, I propose that in describing a field in a transitional stage, it could be important to attend to the cross-disciplinary connections that other researchers have made and how the field might benefit from other disciplinary perspectives. When is it important for the researcher to focus on relationships between his work in education and foundation disciplines (ex: anthropology, computer science, etc.), and to what extent it is important for the dissertation literature review to extend into fields that contribute to the academic area? While Boote and Biele (2005) emphasized the importance of international literature, it would seem that neighboring disciplines present a similar opportunity for emerging scholars to look for comparisons.

---

[2] See the Geologic Society of America Data Repository (www.geosociety.org/pubs/ft2007.htm) as an example.

Overview of This Literature Review

This literature review is situated at an important time in the history of research into how information is used in American education. The majority of the research that I am reviewing in detail has appeared only within the last few years and shows evidence of an intellectual area with boundaries that are "very much under construction" (Moss and Piety 2007, p 3). This construction process and how it may proceed are important considerations in this review because so much of the work is recent, and empirical norms for these works appear to be still in the formative stages. Therefore, this review treats the existing literature as both a conceptual frame to indicate the research foundation and contribution and as a source of data. The literature is analyzed with respect to its contents. This analysis supports a depiction of the historical period that the literature and this dissertation study are situated within.

I have three broad goals for this literature review. First, I want to summarize the empirical nature of the literature as it is. Second, I want to use the literature to describe the historical and developmental positions of this domain of inquiry, to push beyond the statement by Moss and Piety (2007) above and further describe the construction process evident in this literature. Third, I want to understand what type of foundation this literature is for my investigation and what else I need to add to it.

The guiding principle in this review is the appropriate representation of different authoritative perspectives on systemic uses of information in education. Rather than a narrative structure that would have me trace through different studies -- the development of topics such as "the role of leadership in information use," or "the emergence of technological infrastructure," I will present clusters of relevant research grouped into broad categories to represent some of the diversity of the available literature. There are a number of reasons for this approach. One is that all of this literature has been published relatively recently, so the evolution of concepts across studies has data points that are often contemporaneous. In addition, since I am using this literature in an evidentiary way to support some claims about the developmental nature of the field, by presenting a series of short research case studies that are then grouped by type I can demonstrate some of the important variations and similarities in this literature more effectively. Further, while common themes are emerging across this varied corpus, some of the variation in the literature makes building a case upon these themes problematic without considering whether the authors are in fact reporting the same organizational processes when they use the same terms. Finally, this categorical organization also allows me to refrain from introducing an implicit conception of the field as a narrative might. The categorical organization will help as I raise questions at the end of this chapter about where the field may go from here.

Criteria for Inclusion

The review process combined a general literature scan using keyword searches with issue-by-issue scans, as well as bibliographic searches of some recent special publications focusing on the topic of data/information use in schools.[3] Preference was given to peer-reviewed educational journals, as well as selected books and reports from researchers and institutions that publish in these types of journals. A guiding principle in selecting literature for inclusion in this review was that it should be research that describes systemic uses of student performance data. The definition for "systemic" was that it should include more than one organizational component and should involve the types of decisions that could be included in organizational routines (Cohen and Bacdayan, 1994). The key criterion was that the research needed to provide some *authentic description of practice* or be based on a survey that provided some methodologically grounded way to understand actual practice. Publications with recommendations without a substantive empirical component were excluded. My goal has been to ground this review in empirical work since, as I will show throughout this chapter, there is enough important variation in approach among the empirical works without introducing additional layers of complexity from conceptual approaches that have limited connection to current practice. Additionally, literature that primarily addressed practitioners, for example the journal *Educational Leadership*, is not included in this chapter, although I reviewed it in preparation and a content analysis of it might strengthen a more holistic understanding of the development of topics in this new area.

[3] A computer crash in late 2006 destroyed the document that described which journals and which issues I searched and how I searched them. I have considered recreating it later in the writing process.

For a piece of literature to be included in this review, the information had to be relevant to students now in the schools and involve multiple practices. Information or knowledge that came from research (Louis and Dentler, 1988) or was inherent in individuals’ local understanding (Honig, 2006) was generally excluded, even though these topics could occur in the same social/organizational configurations as some of the studies that are included. Also, studies that focused on the consequential use of information that is contained within a single context, such as formative assessment (Black & Wiliam, 1998), were not included in this review, although this kind of literature does enter into the analysis of the dissertation.

As clusters of research were reviewed, qualitative criteria also were used to exclude some publications that did fit the general criteria but where the empirical component was light or presented opaquely, or where an existing process was being studied rather than a plan and hypothetical uses of information. Therefore, important articles such as Thorn’s (2001) *Knowledge Management for Educational Information Systems: What Is the State of the Field?* was not included, because it focuses on the planning and development of operational data systems rather than showing in detail their use in practice. Similarly, while two books that describe how data can be used in educational activities are included, others such as Crieghton’s *Schools and Data: The Educator's Guide for Using Data to Improve Decision Making* (Creighton, 2001) was not included because the empirical component was not a major thrust of the publication.

In the interests of space, while I did review several international publications, I have not included them in this review. as their research was somewhat distant from the conditions of educational practice in the United States. Also, in the interests of space and in keeping with the categorical structure of the argument, I did not include a few articles that could have been included if my goal was to produce an exhaustive list, rather than a representative one. When given a choice of sources, I gave preference to the authority of peer-reviewed publications. I also do not include any literature that focuses on, nor do I highlight the aspects of literature that I do review, discussing the differences between information, data, and knowledge. While this is an important conceptual issue (see Phillips, 2007 for an overview), it is beyond the scope of the current study.[4]

Organization of the Review

I have organized the review into three sections, with a concluding section that helps identify contributions this dissertation can make. First is a section titled *Early Work from a Systemic Perspective*, which is a very brief acknowledgement of some movements that could be identified as precursors to the current wave of research that will be featured in the body of this chapter. I include this introductory discussion to historicize the more in-depth analyses that follow. The next section, titled *Recent Empirical Work*, includes four clusters of contemporary research that explore aspects of information in schools. These four groups are:

- Studies/approaches to information use in education from an organizational perspective;
- Studies of information use in education from a technology or artifactual perspective;
- Problem-focused literature where information use appears in an important way; and
- Conceptual frameworks about information use in education based on empirical results.

This classification system allows for the diversity of this literature base to be explored and compared. Interspersed in this sequence of reviews are reflective comment sections that provide a running commentary on what I believe these studies are showing.

Before concluding the literature review, I first reflect on why it is that this literature base presents a conundrum in terms of using it as the basis for specific research questions. My argument is that while there are important elements of coherence across this work, it does not yet fit into what might be what Kuhn (1970) called a *normal science*. In a normal science, according to Kuhn's view of scientific paradigms,

---

[4] As a general statement, I adopt a position that the social factor of usage in quantitative information is essentially the same kind of process as the use of language and the process of creating and using texts – a process Iedema discusses as *resemiotization* (Iedema, 2003) – and that would be studied from a linguistic frame.

questions are framed by common theoretical models and a body of literature that is similarly based on the common paradigm.

Finally, I present a concluding section that performs several content analyses across these studies to show characteristics of this body as a group of possibly pioneering studies into a new area. I discuss their intertextuality through a bibliographic analysis in order to assess the possible historical significance of this work.

Some Early Work from a Systemic Perspective

When considering the systemic dimensions of educational information use, it is important to note that the current literature does not represent the beginning of scholarly work in this area. While space does not allow a thorough treatment of the precedents, I offer the following vignettes as a modest historical treatment and to reinforce the position that the recent wave of studies that I will review below did not appear from thin air or only in response to NCLB, but can be seen as historically related to other work that preceded and overlaps with the current research into evidence use. Each of these areas has a systemic implication but from different conceptual ecosocial levels (Lemke, 2000).

Measurement-Driven Instruction

One of the earliest appearances in the literature of the relationship between assessment and instruction where the assessment information was given a consequential role in instruction was the measurement-driven instruction discussion often associated with the work of Popham (1987) and Airasian (1988). This work is situated mostly at the level of the classroom, and at the time this concept emerged, educational assessment had been considered something that was born of instruction rather than having a role in directing it. Concomitant with the development of standards based-reform, Popham advanced the proposition that testing could be used to drive instructional decisions. Measurement-driven instruction can, then, be considered as systemic primarily at the classroom level. Airasian (1988) contributed a complicating perspective, indicating that while it is possible for tests to productively drive instruction, there is a range of other circumstances that may cross purposes with that goal.

The Educational Indicator Movement

During the late 1980s and early 1990s, during the first Bush administration, a federal policy initiative culled together a team from inside and outside of education to propose an indicator system for education (Bryk & Hermanson, 1993). The panel, including leaders from business and technology, recommended indicators in six categories, of which student performance was just one. These indicators can be considered systemic on a national level. In addition to testing (called learner outcomes), the panel recommended indicator categories (with subcategories) for quality of educational institutions, readiness for school, social support systems, economic productivity, and equity. When one considers the impact that NCLB has on education in the current period and its emphasis on test scores, which are only one facet of the recommended indicator system, it is interesting to reflect upon what NCLB would have been like if it had been sensitive to this broader range of information. Rather than simple instruments of compliance, the panel recommended that indicators be seen not as instruments wholly owned by policymakers, but as mechanisms to enlighten constituents about the educational process. While the literature I selected in this literature review prioritizes student performance indicators, it is worth noting that many of the other categories of information described in this early work on indicators are also present in many of the sources I analyzed.

Systems View of Educational Testing

Frederickson and Collins (1989) produced an important and often-cited discussion of validity by placing educational testing into a systemic frame. They position their work as an alternative to what they call *passive indicators*. They also construct their argument more narrowly than the educational indicator movement. The test, according to them, should be seen in terms of the types of durable effects it has on a range of activities, such as curriculum and instruction. The test is not only then a measurement device or an instrument for ensuring accountability, but a participant in an organizational ecology and should be evaluated in terms of its systemic consequences:

> The test scores, rather than playing the role of passive indicator variables for the state of the system, become the currency of feedback within an adapting educational system… A systemically valid test is one that induces in the education system curricular and instructional changes that foster the development of the cognitive skills that the test is designed to measure. (Ibid, p. 27.)

By focusing on feedback and the adaptive nature of the educational system, Frederickson and Collins were hypothesizing a potential for organizational cohesion and responsiveness. While others, including Cohen (1995) and those promoting the discussion of educational systems as loosely coupled systems (for example: Orton & Weick, 1990; Spillane & Burch, 2004), raise questions about the ability of educational organizations to act in this way, there are indications in some of the recent empirical work that it is at least *possible* for educational organizations to use test information in responsive ways, although substantive questions remain about how this happens and in which contexts, and how researchers can produce durable warrants about practices that span many activity structures. The perspective advanced by Frederiksen and Collins (1989) can be considered systemic on the local school operational level.

Recent Empirical Work

This recent empirical work is organized into four categories. The first is a set of studies that present an *organizational* perspective. They discuss information/data use as an aspect of what a school or school system does. In these studies, the technology and information are important, but not the main focus. Studies in the second group have in common a technological or *artifactual* perspective. In these articles, the emphasis is on a specific technology or tools, such as a data warehouse system or user interface, that are being used in an educational setting. While there is a good deal of overlap between these literatures, their separate presentation helps in making important points about the evidence base. The third group involves *problem-oriented* studies, where the use of data is important for the study, but not the overall research perspective. The fourth group includes four publications that attempt to help build *explanatory frameworks* to describe some of the fundamental processes at work across educational organizations. These framework pieces are also based on research and raise general questions about when and why information would be used for different purposes. Interspersed in the presentation of these groups are reflective sections that make observations about the nature of the lessons I think these researchers have to teach. This approach is an attempt to mitigate the length of the list-like structure and to develop common themes while maintaining the separate character of the different work.

At the beginning of each section of studies I present a table that summarizes the literature that will be discussed in the section. These tables can be considered as parts of a single table that has been divided into several parts for readability. In these tables, one citation is used to index work that may in some cases appear in different forms. When there are several sources for the research, I will refer to other relevant publications in the discussion that is identified by the single identifiable reference or source. The second column in these tables provides a brief description of what types of information the study is concerned with. This may be helpful for readers to see how this varies across studies in all four groups. The third column is called *systemic aperture.* This is an approximate way to show the parts of the educational system that are being considered by the researchers and highlights some important differences in context that occur across this body of work. The fourth column summarizes the research approach presented in the source, which I treat in greater detail in the conclusion of this chapter.

Organizational Perspectives of Information Use in Education

This first group of studies includes two books written by academics at research universities; one report developed by university researchers at a nationally-recognized policy research consortium; a journal article by an educational research firm that synthesizes information available in a book about a special study, and a research article based on a dissertation. The books and report have many properties of peered-reviewed research literature, including references to theories of organizational issues in education and to contemporary studies into the use of data in schools. The books have been reissued in repeat printings and the report is well cited in the academic literature. While they all focus on different aspects of education, these publications (Table 2.1) often emphasize the possibilities inherent in information use in education over the problems associated with its use.

Table 2.1 - Literature from an organizational perspective

Source	Types of Information	Systemic Aperture	Research Approach
a. Using data to make better educational decisions (Streifer, 2002)			
	Standardized test scores, grades, public data for a range of subjects	Districts, schools	Personal experience narratives, informal cases
b. Data-Wise (Boudett, City, & Murnane, 2005)			
	Multiple sources focusing on literacy and math	Schools, school groups, teachers	Personal experience narratives using two informal cases
c. RAND's evaluation of the Institute for Learning (IFL) system (Kerr, Marsh, Ikemoto, Darilek, & Barney, 2006)			
	Mostly external state and district tests for literacy and math. Learning Walks SM	Districts and schools	Case studies within larger program evaluation research and survey
d. America's Choice selected schools (Supovitz & Klein, 2003)			
	Three broad categories of internal, external, and school-wide. Mostly literacy and math	Schools, school groups, and teachers	Several related cases selected for exemplary data use and research convenience and a survey
e. District influence on teacher use and conceptions (Young, 2006)			
	Literacy in upper elementary school	Teachers within districts	Embedded case study

<u>Using Data to Make Better Educational Decisions (a)</u>

In *Using Data to Make Better Educational Decisions*, Streifer (2002) draws upon his experience as a school administrator in several different districts to show how a school system can benefit from the use of information related to student performance. Of the various publications reviewed here, this one is among the closest to what might be considered a program evaluation piece, because the decisions that are discussed are at the level of a school system. Streifer uses an informal case-study narrative approach featuring three schools: a K-8 school, a middle school and a high-school, and shows how a range of information including the state performance test, a ranking of instructional materials difficulty, student ranking against state measures, and grades can be used to identify instructional issues. Using his experience as an administrator, he places information use into a systemic context so that the reader sees why it is important to get new information or use the information at hand to address some basic educational system concerns. In one of his analyses, district administrators used the information to diagnose systemic breakdowns and set performance targets for English language arts. In another, the district administration identified possible interventions for middle school mathematics. In two other analyses, he uses a university study of school data to identify possible problems with textbooks and possible grade inflation in high school science. Streifer includes several discussions that look at information longitudinally. The Connecticut state mastery test is used in three of the four analyses he presents.

Streifer makes a number of recommendations that are also found in other literature, including the importance of visual displays of information and the need to make data accessible to busy practitioners.   He recommends visual devices as a way to structure a school's data analysis process; using concept maps (Novak & Gowin, 1984) to help practitioners develop organized questions where data can play a role. He pays particular attention to issues of technological infrastructure, describing relational databases and then data warehouse technology and data mining. He frames his approach within a cultural and systems model where these intangibles are critical success factors, rather than a model where the technology provides ready answers, as some other may have done. He describes cyclical and iterative developments of organizational capacity to produce and use information.

Data Wise, a Harvard Project (b)

In *Data Wise: A Step-by-Step Guide to Using Assessment Results to Improve Teaching and Learning* (Boudett, City, & Murnane, 2005), a number of faculty, researchers, and students from Harvard Graduate School of Education propose a school-centered approach to data use. Where Streifer provides a district level perspective, the focus in Data Wise is inside the building, with discussions of how school leadership (both school building leaders and data teams) and individual teachers can use data to identify and analyze learning problems that may affect all or a subgroup of students. This project draws off of a partnership with the Boston Public Schools and features two informal case studies, an elementary school and a high-school. It provides examples of how a school's professional community can use information activities to raise scores, develop common understandings, and plan their school's course of action.

Similar to Streifer (2002), Data Wise also emphasizes a holistic and cultural approach to data use, recommending an eight-step iterative process[5] that begins with a data inventory and then continues to refinements of data inquiry and action based upon that inquiry. The data inventory described for the case-study primary school resulted in nine different forms of evidence:

1. State skill assessment (reading, ELA, math)
2. Reading: observation survey (Reading), developmental reading assessment (DRA)
3. Stanford 9
4. English proficiency
5. Reading checklists
6. Running records (Clay, 2000)
7. Writing samples
8. District math assessment
9. Unit assessments (Math)

The book goes on to describe how these partner schools used data. In the primary school, the leadership and data teams used the data for comparisons, trends, and identification of problems in practice to raise student scores. They found the conversations around data supported common understandings among teachers in the definition of learning problems and performance expectations, as well as more standardized grading and definitions of best practices. At the high school, the principal and the math department combined an item analysis of state comprehensive tests results and student work to diagnose instructional problems. Their triangulation of these data sources allowed them to see why students might be struggling with multistep math problems. They then used video from the Third International Mathematics and Science Study (TIMSS) to show alternative practice. The authors also explore data being used in planning processes.

Much as Streifer did, the DataWise authors discuss the importance of visual displays of quantitative information (Tufte, 2001). Where Streifer (2002) recommended

[5] This process bears significant similarity to the process described by Love (2002).

using concept maps to frame problems, DataWise proposes some sample paper forms to help guide the practitioners' analyses.[6]

Reflection: the Issue of Comparability

These two studies introduce some of the ranges of application of student-related information to practices of making educational decisions. They also help to introduce a question that will carry through the other examples that follow: In what ways are these sources part of the same literature or the same field? While they both involve the use of information in an educational context, their systemic aperture is significantly different. They would occur in different organizational, and in most cases physical locations, with one being centered in a district office and the other in a school. While Streifer (2002) does include school-level analyses, these were activities conducted by university researchers and graduate students and did not actually involve school level personnel in the way discussed by the Data Wise team, where the school personnel took ownership of their instructional information. In addition to the issue of the different systemic apertures between these publications, there are also issues of the purposes for and the nature of the research. The Data Wise book focuses on teachers and collegial groups inside the building, while the other focuses more on accountability from outside the school building.

This is an issue that will appear again in some of the reviews that follow, because neither study provides the kinds of details that would be found in a report disseminated through a scientific journal.

RAND's Evaluation of the Institute for Learning (IFL) System (c)

Strategies to Promote Data Use for Instructional Improvement: Actions, Outcomes, and Lessons from Three Urban Districts (Kerr, Marsh, Ikemoto, Darilek, & Barney, 2006) is from a comprehensive program evaluation of the University of Pittsburgh's Institute for Learning (IFL) school reform program. Researchers at RAND

[6] My professional experience includes the use of various organizational analysis methods with similar visualization and organizing devices. While across different methods there were different forms of work documents, they tended to serve similar functions of providing a structure or scaffolding for the group analysis activity.

conducted this study and focused on three urban districts (see also Marsh, Kerr, Ikemoto, Darilek, Suttorp, Zimmer, & Barney, 2005). The researchers report that two of the three districts engaged in data-driven decision-making to a much greater extent than the third and largest one (approximately 80,000 students compared with 30,000 for each of the other two). Further, the two that did invest heavily in data for instructional purposes did so in different ways. This research discusses information from external sources (district and the state assessment), interim assessments (six-week tests in one district), and classrooms, although the classroom assessments are only mentioned in terms of some teacher preferences for them. And, similar to Streifer (2002), this is a study focused on the district level.

Of the two districts that had an emphasis on data use, one used the school improvement process as a vehicle for change while the other focused on a data system technology and the development of interim assessments that were linked to this system. For the school improvement program, the district encouraged the use of state test results and a structured program to allow them to set a narrow set of instructional goals. The survey results indicate that school-level personnel involved in the school improvement process and professional development planning found the data to be an effective tool. At the other district, the interim assessments system proved popular for school and district leadership, while less so with teachers. Some teachers also used item-level analyses to focus on specific student learning problems, while many reported that their own classroom assessments were more useful for planning purposes. This district also indicated that technological problems in terms of data access and information integration were important factors affecting the success of the program.

In addressing the general issues of data use in school, these authors point to the importance of accountability policy in understanding practice. They also highlight the importance of timely information and indicate that one of the reasons the third and largest district did not achieve a strong systemic commitment to data use was related to difficulties in information access. Capacity in terms of staff and technological infrastructure were also highlighted, and the authors further discussed the perceptions of validity and usefulness on the parts of the practitioners.

America's Choice Selected Schools (d)

The next example of research into data use comes from a study of a select number of schools implementing the America's Choice school improvement program. *Mapping a Course for Improved Student Learning: How Innovative Schools Systematically Use Student Performance Data to Guide Improvement* (Supovitz & Klein, 2003) describes a study of schools nominated by the school system's program leaders based upon exemplary use of data. The authors state clearly that their research data are not representative. Not only are the schools selected, but the teachers that were interviewed were nominated by their school principals and so represent probably outliers in terms of their practices. For these schools the researchers performed site visits and interviews. Additionally, they used a survey sent to 68 school principals randomly selected from schools that had completed their first year of the program.

Table 2.2 - Data types described in Supovitz and Klein (2003)

Classroom Data	School Data	External Data
Running records, writing notebooks, conference logs, source books, math journals. Data provided by leaders, student-generated physical artifacts, other innovative teacher-developed sources.	Running records, theme assessments common to the school	State assessments, district assessments, New Standards Reference Exam (NSRE)

Supovitz and Klein define categories of data that were used in America's Choice schools: classroom data, school data, and external data, as shown in Table 2.2. While the classroom data might correspond to types of assessments often called formative, and the external data would include categories of assessments often referred to as summative, the middle category of school data represents a relatively new type of assessment process and one that deserves attention because it can be managed within the school or district. These are local assessment systems and can be helpful in setting common standards for teachers teaching the same classes and also would typically involve school leaders. In other research, these types of systems have been called *interim assessments* because, even when the underlying metric is the same as some other measurement system (ex: running records), the use at the school level is less frequent than in classrooms but more frequent than external assessments.

The uses of this information range from individual planning for targeted students by principals to professional development planning, setting goals and targets and internal school communication (celebrations, visually displaying progress, etc.) by leadership teams that include principals and coaches. Teachers reported that these data were used to align the topics they taught and the emphasis they gave within those topics, as well as communication with parents. One of the report's major themes is that available data are often not used:

> There are potentially both a rich array of data available to teachers and leaders within their schools and a multitude of ways in which those data can be used for instructional guidance, institutional planning, organizational support, and cultural influence. Most of the data available to schools goes unexamined. Yet, data can provide a solid bedrock from which to base well-considered courses of action and to test whether past decisions have paid off. (Ibid, 2003, p.39.)

The authors make recommendations of greater use of school-wide data systems that integrate instructional/classroom, external and school-wide performance information. They provide models for how these different data sources can be used together in developing a culture of inquiry within the school.

District Influence on Teacher Use and Conceptions (e)

Young's *Teachers' Use of Data: Loose Coupling, Agenda Setting, and Team Norms* (Young, 2006) reports on four schools in two districts. Using an embedded systems framework to build a nested case study deign, Young describes some factors that affect how teachers collaborate, using data from both district and vendor developed sources. Her research sites are urban schools and her focus is on decisions related to students in the upper primary grades. Young's work is grouped into this first category of publications because it does focus on organizational aspects of information use; and its systemic aperture on teachers within districts complements the first four examples. This work could also have been placed into the third category. problem- focused research. because it is a tightly-focused study, while the first four studies presented in this group tend to be broader and less specific in terms of research questions.

The four schools in two districts include diverse community and student types (ex: poor to working class). Young shows a range of data-related functions in which teachers participate, including reporting, providing resources for instruction and supporting team meetings for school management collaboration. For example, in one of her schools: "…the second-/third-grade team primarily used assessment data and their observations of student work and behavior to group students for literacy instruction, to move students between groups midyear, and to create and review intervention strategies for individuals." (Ibid, p.526.)

Her study shows that all levels, from the district staff to teachers, affect how data is used. The information profiled in this study comes from different sources, some developed in the district and some developed by vendors. This study shows some of the variations that may be expected when practice is studied closely.

Reflection: Optimism and Reality

These five examples from the literature provide some of the most comprehensive examples of how data can be used systemically and instructionally within educational contexts. They describe a range of situations relevant to different grades and subjects, although with stronger representation of elementary math and literacy than other areas. These studies rarely describe typical schools or schools selected at random. And care is needed in extending what is reported in these studies to a wider conception of educational practice today. Still, these reports include some potential heralds of future practice. The interim/school-wide assessments from the second two studies are an example of an important elaboration on a traditional formative-summative model, with information that relates to a school or district's students being identified as relevant for school-level leadership and collegial collaboration. In fact, these studies indicate that the mediational power of student information is pushing on what has historically been an individual and private practice (Jackson, 1968).

Artifactual Perspectives of Information Used in Education

This next group of literature, summarized in Table 2.3, looks at the issue of information in schooling from the perspective of artifacts and technologies. The types of information, data storage systems, and access methods these authors discuss also appeared in the literature reviewed above. But in these studies, the artifacts move onto center stage by being used in the framing of the research and as a focus of the publication.

Table 2.3 - Research from an artifactual perspective

Source	Types of Information	Systemic Aperture	Research Approach
f. The Annenberg "From Data to Decisions" report (Meiles & Foley, 2005)			
	Data warehouses with range of performance and non-performance	District, schools, teachers	Informal case-studies with quotes.
g. The Grow Network study (Brunner, et al. 2005)			
	User interface with scores from city-approved standardized tests	Districts, schools, teachers	Three phases: 1) district, 2) school ethnography, 3) survey
h. Full school faculty use of data (Wayman & Stringfield, 2006)			
	Data warehouses and supplemental data systems with performance information	Schools	Informal case-study with short site visits
i. Quality School Portfolio (QSP) (Chen, Heritage, & Lee, 2005)			
	User interface with range of performance, demographic, and historical data	Schools	Evaluation study using telephone interviews

The Annenberg "From Data to Decisions" Report (f)

In *From Data to Decisions: Lessons from School Districts Using Data Warehousing* (Meiles & Foley, 2005). eight large districts in different states were studied to understand how they were using data warehouse technology. The authors report on lessons learned in support of similar knowledge management efforts in other districts. As with the studies from the last group (and many that follow in this analysis), the sites selected are not representative. They are districts identified as pioneering, "three of which had won national recognition for using technology effectively." (Ibid, p. 3.) They range in size from over 13,000 to more than 240,000 students. These would tend to be districts large enough to have some infrastructure to be able to implement advanced technology.

The report does not detail its methods other than to indicate that interviews were used, even though the researchers are from a school reform institute housed at a major university (http://www.annenberginstitute.org/). Rather, it begins by describing the common condition that many schools find themselves in, with an abundance of data that is scattered in inaccessible and antiquated technology, often controlled through a small number of individuals who have the capacity to access the data. They then describe data warehouses in general terms as the integration point for information and data that can provide broad and flexible access.

The information that is in these different systems is also only described in general terms and rarely by specific instrument or measurement approach. In addition to performance data, student attendance, after-school activities, and discipline and behavior information were reported as being housed in data warehouses. Examples of teachers, principals, coaches, and administrators using this information were also provided in short case vignettes. The information-related activities briefly discussed in this report include teachers identifying learning gaps; discussions around performance standards; inquiry-based analysis using multiple data sources to understand teacher work from afar, and general efforts to improve scores – all of which are consistent with the literature reviewed previously.

Much as some of the previous studies have indicated, leadership is key to implementing these data warehouse solutions, and the process of implementation is iterative rather than completed in one event. The format and usability of information is stressed, as practitioners should be able to find the information they need as quickly as possible. This emphasis on making the information easy to access is included in the development of customized reporting options tailored to specific practitioners. In addition to the technology and leadership issues, these authors stress (as others have) the importance of what Petrides and Nodine (2003) call an "information culture."

The Grow Network Study (g)

Linking Data and Learning: The Grow Network Study (Brunner, et. al., 2005) evaluates a way of presenting assessment information in the New York Public Schools.[7] The Grow Network is a subsidiary of McGraw Hill Publishing, and it can be thought of as a study presenting information across multiple levels of the system. For the most part, the Grow Network was based upon standardized test data approved by the city. The data was from math and literacy and the interface was "…a mix of print- and Web-based reporting systems." (Ibid, p.242.) The research design had three distinct stages: a focus on the central office, ethnography in 15 schools within four districts, and then a system-wide survey. The research was conducted during a period of transition for the school system and went through a difficult phase where the schools seemed to be responding to another set of top-down initiatives, so the evidence of practice the researchers intended to document was elusive.

> Not only did sweeping administrative and instructional changes take place, but the new leadership also introduced another accountability resource, supplied by The Princeton Review, that provided teachers and administrators with testing results based on assessments administered three times during the current academic year. (p.248)

Despite the introduction of an alternative information system with intermediate assessment points that interrupted the research, the study showed a number of ways that information was useful in organizational activities. For example, school leaders (considered in this study to include school and district non-instructional staff) used the system to identify gaps, target resources and plan future programs. Teachers reported using the data to identify areas of weakness, tailor instruction and individual education plans (IEPs) and focus on "bubble kids" (Confrey & Makar, 2004) - those near the upper boundary of a proficiency range who with some help could reach the next level. Teachers also grouped kids and assigned performance targets (cut scores). Both teachers and administrators reported the data supported conversations with parents, students,

---

[7] There is also a report available from the Consortium on Chicago Schools Research that provides some description of Grow Network in Chicago Public Schools (http://ccsr.uchicago.edu/publications/p66.pdf)

fellow teachers, and administrators about student learning. The details of these conversations were detailed in the study. The system was also used to support planning for teacher's professional development, and teachers reported that the information was helpful in shaping their practice by allowing teachers to reflect upon their own work.

Full School Faculty Use of Data (h)

In *Technology-Supported Involvement of Entire Faculties in Examination of Student Data for Instructional Improvement*, Wayman and Stringfield (2006) also study data warehousing, but focus on schools rather than districts. They looked at three schools in three different districts. Their study was conducted with brief site visits for two of the schools and telephone interviews for the third to understand the schools' implementation of commercial data warehouse technology, as described in Table 2.4 below. Each of the schools in this study was using a combination of a central data warehouse and some other information source, either from the Northwest Evaluation Association (NWEA) or from a vendor. Like the districts discussed in the "From Data to Decisions" report, these schools were maintaining more than performance scores. And, like many of the reports presented thus far, math and literacy were the focus of the data use.

The participants in this study reported that they use these information systems for planning, teacher data analysis, profiling of student needs, and reflection on practice. Some common themes that emerged, consistent with other studies, were that principal leadership was key to the success of the system and that it is important to consider teachers career positions and use non-threatening triangulation of data sources. Like other researchers, Wayman and Stringfield report that user interface and presentation is important for the success of the data initiative, that information needs to be timely, and that practitioners need adequate time to be able to process it. They also report that using data well often requires a lot of data for the analysis to be productive and that technology interoperability and personnel capabilities/requirements presented challenges for the schools. While there was a sense that the authors found practices that involved data did improve school professionalism and teacher efficiency, they reported difficulties for teachers to actually connect data to practice.

Table 2.4 - Information types in Wayman and Stringfield (2006) study

School	Type of Data Warehouse	Other System
Small Northeast PreK-5	EdSmart Data Warehouse Reading & writing, cognitive skills data	NWEA Reading, Other (unspecified)
Large grade 5-6 school in small suburban district in the northern U.S.	SchoolNet Data Warehouse Reading, cognitive skills, district developed assessments, student demographics, status	Renaissance Learning Reading and math
Grades 6-8 in large district in southern U.S.	SchoolNet Data Warehouse Reading, cognitive skills, district developed assessments, student demographics, status, special ed	Renaissance Learning Reading and math

The Quality School Portfolio (i)

Identifying and Monitoring Students' Learning Needs With Technology (Chen, Heritage, & Lee, 2005) discusses a software system called Quality School Portfolio (QSP). QSP is a set of software tools, under development since the late 1990s. It has a web interface that can be used to present educators (administrators and teachers) with a variety of information that includes assessments, demographic data, groupings of students, student home information and student personal interests, and student academic history. Core features of QSP are a longitudinal student database, disaggregation capabilities, report functions, goals and monitoring functions, a gradebook, and a digital portfolio. Color-coded tabs organize all of the QSP functions, akin to a filing system. (Ibid, p. 312.)

Like the Grow Network, QSP is primarily a way to access and integrate information that is captured elsewhere, although it does have a central database. Unlike the Grow Network study, the schools using this tool (at least 22 across the nation indicated in the study) are able to decide which data sources are used within it. The interface has many dashboard-like features and allows grades to be represented in a score fashion as well as by standards. The evaluation of this research tool is ongoing, and similar to other reports of information use, the QSP evaluation study tells us that the use of QSP allows practitioners to identify students' needs, promotes collaboration, allows for greater inquiry into and reflection upon practice, and can be an aid in planning.

Reflection: Emergent Themes

In looking across the two approaches to studying information using organizational and artifactual perspectives, there are some broad topics that appeared in both sets of sources. I have grouped these broad topics into three categories, as shown in Table 2.5. The first category represents the types of critical factors that are given by authors as reasons for a data initiative to succeed. The ways that different authors present these topics often varies. But they are all presented as essential to the success of the data effort. The second category involves how information is put to use: how the researchers and their participants represent the ways that information availability relates to practice. In the third category are topics that seem to address a more sophisticated systemic use of information in some way. One of these advanced features is interim assessments. Another is the use of data over time/repeated measures, also called longitudinal data. A third addresses the integration of multiple data sources into some complex data-informed decision-making practice.

This intermediate analysis suggests that despite some important variation in studies and many studies using alternative research genres in terms of study details, there are some common aspects that arise when the issue of the use of data and information in the schools is studied. These initial findings suggest a set of core issues such as reflexivity, collaboration, and leadership where information of many kinds, particularly student performance information, can be instrumental. The emergence of similar patterns of topics across both sets of studies should be received critically. While these are broad topics -- and the ways that concepts such as communication and leadership are discussed across the different research -- the literature may differ in important ways. Taken as indicators of similar broad categories from a very small number of studies, they also indicate interrelationships of organization and artifact. They indicate that this issue of artifacts in organizations may be productively approached by looking either at the organization and its artifactual traversals or at the artifact (a text) and the paths it takes to arrive at an understanding of this process.

Table 2.5 - Summary of broad topics in organizational and artifactual focused literatures

Broad Topic	a	b	c	d	e	f	g	h	i	
	Organizational					***Artifactual***				
1. Importance of leadership/org. culture	X	X	X	X[1]	X	X		X	X[2]	***Critical Factors***
2. Technological capacity and infrastructure	X	X		X		X	X	X		
3. Capacity process: training, learning	X	X	X			X		X		
4. Visual presentation/timeliness	X	X	X	X		X	X	X	X	
5. Reflection/communication/collaboration		X	X			X	X	X	X	***Purposes***
6. Planning (ex: programs or PD)	X	X	X	X	X	X	X	X	X	
7. Grouping of students/ targeting instruction			X	X	X		X	X	X	
8. Interim assessments (school or district)			X	X	X				X	***Advanced***
9. Longitudinal data	X	X				X			X	
10. Use of multiple data sources	X	X	X[3]		X		X	X	X	

(Header: **- - - Reference - - -**)

1 - In program, but not tied to the data.
2 - Not mentioned in these terms
3 – Multiple forms of assessment data

In turning to the next group of studies, studies where information use also appears significantly but is not the focus of the research, these broad topics/themes will continue to be manifest. But, the practices of research and reporting will be more standardized and the findings often more focused and nuanced.

Data Use in Problem-Focused Literature

This set of literature is the largest group I am reviewing and perhaps the most transparent regarding research practices. All but one publication are from leading peer-reviewed educational research journals. Unlike the two groups discussed above, information and data are not general concepts applied to educational situations but are treated specifically within specific problem analyses. In these studies, identified in Table 2.6, student performance is the most common type of information being considered.

Table 2.6 - Problem focused studies where information use is considered

Source	Types of Information	Systemic Aperture	Research Approach
j. How teachers make sense of policy (Coburn, 2005b)			
	Literacy assessments in elementary schools	Teachers within schools	Formal case study
k. Mechanism for connecting teachers and the system (Coburn, 2004)			
	Range of literacy assessments (district, school, commercial)	Teachers within schools	Formal case study
l. Schools responding to accountability(Diamond & Spillane, 2004)			
	Math and literacy scores from Iowa Test of Basic Skills	School leadership	Informal case study
m. Greater teamwork with intermediate assessments (Halverson, 2003)			
	ITBS, State tests Interim tests for math	School leadership	Case study of one school
n. Science from the school to the statehouse (Falk & Drayton, 2004)			
	State science test Classroom practices	State, district, school, teachers, curriculum coordinators	Informal case study
o. Looking at teachers understanding of data (Confrey & Makar, 2004)			
	State test and TIMMS	School and teachers	Case study of one school

How Teachers Make Sense of Policy (j)

The role of non-system actors in the relationship between policy and practice: The case of reading instruction in California (Coburn, 2005b) explores how teachers mediate policy. It allows a focused look at the issues of leadership and planning raised in the previous group of literature. Drawing on research into literacy instruction in urban elementary schools, Coburn shows different ways that teachers collectively interpret policies and use assessments both as opportunities to reflect upon practice in meetings and in making instructionally relevant decisions. Teachers' mediation of policy messages occurs, according to Coburn, in informal as well as formal meetings and is not driven by the content of the information alone, but rather by their predispositions and personal histories with previous encounters in their instructional environment (Coburn, 2001, 2004). These teachers are also influenced indirectly by their principal's belief systems based on similar cultural and historical predispositions (Coburn, 2005a).

Using both sensemaking (Weick, 1995) and organizational coupling (Weick, 1976) as frameworks, Coburn positions the teacher as an actor within an institutional context that sends the teacher messages that teachers respond to. Different teachers interpreted the idea of using assessments to inform instruction in different ways. For some "… assessment to inform instruction meant knowing where in the sequence a child was and planning lessons accordingly." (p.153). Others took a different approach, not placing kids in a sequence but rather looking at students' needs individually and relating the information to a predetermined sequence in the same way. In this research, the composition of the teacher teams was also important to the different approaches that teachers took, so that to some extent the approach for a given teacher is individual and to some extent influenced by different institutional groupings.

Mechanism for Connecting Teachers and the System (k)

Another important study by Coburn is *Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom* (Coburn, 2004), where assessment information appears in her analysis of how teachers are connected to their institutional context through different types of coupling mechanisms that Coburn divides into *system* and *non-system* elements. Her study defines actors broadly to include individuals such as colleagues; organizational activities that include professional development programs; and artifacts such as textbooks and assessments. This study shows the artifacts acting as *organizational connectors*; as being instrumental in linked (systemic) organizational action. She shows that for the teachers in her study, assessments from the state test and from the textbooks were ways that the teachers were connected within the system or practice and by which the system communicated with them. While these human and artifactual *actants* (Latour, 2005) were included in a whole range of ways that teachers "read" policy, Coburn's study shows some of the roles that assessments fill within systems of practice. Using a set of focused case studies on individual teachers, she writes about one teacher that:

> When Deanna began working with the professional development provider, she started to make instrumental shifts in her practice…By the end of the year, she was assessing all children on a monthly basis and adjusting the

> composition of her reading groups on the basis of the information provided by the assessment. (p. 36)

While she finds some instances where the assessment is implicated in a specific practice, in other cases Coburn's teachers seem to make more symbolic reference to what their institutions ask without substantially altering their practice. One of the most important aspects of this study is how it shows practitioners may be prone to inaccurately report their use of assessment information.

As with some of the earlier studies, she further illustrates that the basis on which these various forms of information may be interpreted by teachers can be traced to their previous messages within the system; how this diachronic view (Coburn, 2004) can be seen both at the system and teacher level, and how district and school personnel are similarly influenced by policy messages they have received in the past (Coburn & Talbert, 2006).

Schools Responding to Accountability (l)

High-Stakes Accountability in Urban Elementary Schools: Challenging or Reproducing Inequality? (Diamond and Spillane, 2004) looks at how four urban schools responded to high-stakes accountability. While much of the literature in the previous sections discussed information use in positive terms as schools worked to fix problems, this study shows another dimension: information being used to further disadvantage students by grouping. This is one of the few reports discussed here where the information use is portrayed negatively.

Two of the schools in the Diamond and Spillane study were in a probationary status and facing further sanctions because of low performance, and two of them were not. The use of assessment data differed greatly with the nature of the school. Those schools not on probation were able to use the student test scores (Iowa Test of Basic Skills or ITBS) to identify weak areas in their curriculums and develop plans to improve in these areas. Those on probation responded by using the results to identify those students who were close to a cut-off and target instruction towards them. The targeting of instruction towards those "bubble kids" (Confrey & Makar, 2004) implies likely leaving those students far below the cutoff, those considered to have low chances of affecting the school's overall results, behind. This is a situation of accountability response increasing inequity for those most disadvantaged. In this setting, the assessment scores are accessories to decisions taken by educators that can leave some students behind. In a more general sense, this study shows how the way that test information is used can be strongly related to the *content of that information*.

This study and Halverson (2003), to be discussed below, are both are from the Distributed Leadership Project (DLP) funded by the National Science Foundation and the Spencer Foundation:

> "designed to analyze the practice of school leadership in urban elementary schools. Building on theories of distributed cognition, the central goal of the project is to make the 'black box' of the practice of school leadership more transparent by revealing and analyzing how leaders think and act to improve instruction in their school. "

Spillane and his colleagues show how school leadership occurs across different types of professionals, in individual steps, and can often involve the creation and use of assessment information (Spillane, Diamond, & Jita, 2003). These researchers theorized that leaders mediate their roles through a range of artifacts from the routine, such as memoranda, to the occasional, such as meeting plans (Spillane, Halvorson, & Diamond, 2003). While not specifically theorizing the results of tests separately from other types of artifacts, these researchers create a theoretical model whereby the testing practices and testing results are used by school leaders to structure sense-making practices similar to the way semiotic artifacts are discussed in cultural approaches to cognition and communication (Vygotsky, 1985; Wertsch, 1998).

Greater Teamwork with Intermediate Assessments (m)

Halverson's (2003) *Systems of Practice: How Leaders Use Artifacts to Create Professional Community in Schools* reports on how school leaders can implement what he refers to as artifacts for their Activity System (Engeström, 1987). In a move that is similar to the framework used in Coburn (2004), he considers assessment systems in the same study that looks at an organizational routine called "the breakfast club" and a school improvement plan. In a move that is similar to the taxonomy produced by Supovitz and

Klein (2003), he theorized artifacts can be classified as inherited (for example, a high stakes test), locally designed (for example, a classroom assessment), or institutionally inherited, such as a district program or perhaps a textbook that is developed outside of the school but enacted and modified locally. He reports: "The Five-Week Assessment provides another angle on the on-going effects of classroom practice through collaboratively developed measures of student achievement." (Ibid, p. 23.) Using a single case study elementary school, he shows how various leadership tools, including this interim assessment process, integrate the work of professionals in the school. He goes on to report that: "The collaborative development and implementation of the Five-Week Assessment provided needed closure among teachers in the system of practice. The Five-Week Assessment also gave school leaders feedback on how new instructional efforts fared in classrooms" (p. 24).

Science From the School to the Statehouse (n)

State testing and inquiry based science: Are they complementary or competing reforms? by Falk and Drayton (2004) asks questions about the relationship between high-stakes tests developed and administered by the state of Massachusetts and the classroom practice of scientific inquiry. Looking at six middle schools in six different Massachusetts districts, this study bears some important similarities to this dissertation in terms of the field it studies and its scope. The study was embedded in a larger study of the effectiveness of the NSF-funded state science initiatives (SSI) and relied upon "observations of teacher practice, interviews conducted with staff at every level of the system from superintendent to classroom teacher, and artifacts collected from the state, district, school, and classroom." (Ibid, p. 354.) The methodology is referred to as a *continuous comparative approach,* and the results are often presented as case studies using the schools -- three from affluent or moderately affluent regions and three that are urban districts described as "challenged" -- as exemplary of the types of issues encountered when trying to reconcile the vision of learning communicated in high stakes tests with the vision of learning articulated by the professional organization for science education. The authors find, much as Young (2006) did, that the response to the state test for teachers was strongly affected by their district; and much as Diamond and Spillane

(2004) did, that the district response varied between those more and less advantaged. Similar to Coburn's (2004) work with elementary literacy instruction, individual histories in terms of the messages that professionals have experienced from the systems do matter in their use.

Within the advantaged (and to some extent the disadvantaged) groups, the districts took different approaches to their use of the state test information. In considering the overall effect of the state test on instruction, the authors find that

> The differences in strategies suggest that the test itself does not dictate change in instruction; rather it is the district's interpretation of and response to the test, in the light of previous pedagogical commitments (where these exist), that strongly influence the reactions of teachers within their classrooms. (Ibid, p. 376.)

The issues that are being discussed in this research are different from the more common themes of raising scores or equity. They involve what is the best way to teach science. Is it by deep exploration and investigation through inquiry methods, or is it by covering a range of topics as represented on the state test? The SSI had been promoting inquiry, and many of the schools had supported that movement. The role of the science content specialists or curriculum coordinators was also illuminated in this study, with the approaches of specialists from two of the moderate districts highlighted:

> Each of these coordinators had a significant impact on their schools. They each brought a sense of coherence to the pedagogical approach that was being used, and to the definition of what was taught within each grade and within each classroom. However, when the MCAS was introduced these two districts took markedly different approaches. (Ibid, p. 377.)

Unlike the elementary literacy coaches represented in other studies, these positions are district-based. When the authors present a comparison that includes a mention of the curriculum coordinators, the examples are both from the more advantaged schools. During the entire discussion of the three challenged schools, a coordinator is only mentioned once.

In this study there is suggestive evidence that the types of local differences in relationship to information use shown at the school level for elementary literacy and math (Diamond and Spillane, 2004) exist in a similar form at the level of the district with middle school science. Unlike the fine-tuning versus targeting of students seen in elementary grades, this study suggests that in some contexts the responses for middle school science may be coarser. Districts either adopt plans to cover the content on the test and deemphasize their approach to inquiry, or commit to inquiry and address the test within the context of their district science priorities.

Looking at Teachers Understanding of Data (o)

Using Dynamic Statistics Software to Critique and Improve Use of Data from High-Stakes Tests (Confrey and Makar, 2004) presents a focused look at the presentation of accountability information. The work relates in important ways to Diamond and Spillane (2004), as it looks specifically at the types of grouping decisions taken by educators who are facing accountability pressures based on the presentation of a small amount of information.

Against a backdrop of a research collaboration shown in retrospect to be producing important learning gains that was disrupted by the state test, combined with the needs of their practitioner collaborators' to respond to their school's accountability status, Confrey and Makar launch a conceptual discussion about some fundamental aspects of the nature of the scores that carry so much weight for the students and faculty. They specifically look at aspects of natural variation and ask whether those interpreting the scores know or can visualize, with the presentations that are provided to them, the differences between real deficits and statistical artifacts. Describing what led them to this investigation they report:

> This case led our team to question the implications and approaches drawn from the disaggregation of data and the *design of the accountability system* in the case of small populations. It made us aware of the neglect of distribution, sampling variation and inferences of statistical difference … (emphasis added, p ??)

The issue of the disaggregated results leads to discussions of the categories that overlay the entire accountability system: not only the subgroups of students that were required for the reporting of these results (ex: African American, Hispanic, White), but also the designation of special education that plays a role in the educators' decisions about which kinds of students receive what kinds of educational resources. The authors' work provides a strong example of the type classification issues that Bowker and Star

(1999) discuss related to health care in an educational context. They show how classifications embedded within the assessment policy and artifacts can be significant in the kinds of choices made by those who interact with the system.

Explanatory Frameworks from Research

The four studies in this section are more general than those presented above. They tend to be over longer time frames and have larger numbers of participants. These works have in common that they are trying to explain how information practices may work in general terms, through the use of research data rather than purely theoretical approaches.

Table 2.7 - Literature in Review Presenting Explanatory Frameworks

Source	Types of Information	Systemic Aperture	Research Approach
p. Early study of state systems (Massel, 2001)			
	Mostly state tests including performance tests	States, districts, schools, teachers	Three year study plus a survey study
q. Ontology of data-driven decision making (Ikemoto & Marsh, 2007)			
	Typology that includes input, process, output, and satisfaction data	Districts, & schools	Secondary analysis of data from two studies
r. Barriers to teachers using data (Ingram, Louis, & Schroeder, 2004)			
	General categories of systematic and non-systematic information	Schools: teachers and leaders	Longitudinal case studies with interviews & site visits
s. Sub-cultures of beliefs developed over time (Coburn & Talbert, 2006)			
	Looks at research and evidence (scores) similarly Elem. literacy focus	District: upper & front-line School: principal & teacher	Part of two year study of districts used interviews and case studies

Early Study of State Systems (p)

Massell's *The Theory and Practice of Using Data to Build Capacity: State and Local Strategies and their Effects* (Massell, 2001) is from a combination of parts of studies in a large systemic reform research project. It is one of the earliest examples of the attention that educational research has paid to how information from accountability systems was being actually used. This study began with eight states and proceeded in stages, first with interviews and then with visits with district and school leaders. This research was combined with a survey of teachers from the same research program to produce the report that discusses how different states and different districts use mostly assessment data.

Consistent with other studies, this research shows state accountability data being used for a variety of purposes, including alignment of curriculum and planning for school improvement and professional development. The study also shows some districts were encouraging schools to develop local data systems. As with some of the case studies presented earlier, the study showed that the variation in data use was not entirely a function of accountability pressures on schools, but also of the district and school leader's beliefs about the validity of the accountability goals and their belief in the concept of using data to inform educational decisions. This study clearly shows that many of the topics that some of the contemporary literature report as being driven by NCLB actually preceded the legislation and likely have more complex roots.

Ontology of Data-Driven Decision Making (q)

In *Cutting Through the "Data Driven" Mantra: Different Conceptions of Data-Driven Decision-Making* (Ikemoto & Marsh, 2007), researchers analyzed the data sets from two studies, one of which was reported in Kerr, et al. (2006) and Marsh, et al. (2005). The researchers combined an analysis of project documents and survey responses for an indication of their use of data in decision-making. Using Mandinach, Honey and Light's (2006) framework for data-driven processes to describe how data becomes information before being used as knowledge across multiple levels of the system, they developed a taxonomy where both data and the analysis process can be located along separate continua of simple and complex. Joining these two continua into a quadrant, they developed a model that they then used to categorize short cases of data-oriented activity from their larger research. Most of these cases, they found, can be placed in the simple-simple quadrant, indicating that the data is simple (ex: a school's score on a standardized test) and the decision is direct (institute professional development in response). More complex forms of data might include multiple measures or data that required sustained inquiry and complex decisions such as school improvement plans – both examples of the advanced features shown in Table 2.5.

Like some of Coburn's (2005b) work, the researchers alert us to the potential that participant reports should be considered carefully for credibility. They begin with their respondents' reported data usage:

> Educators across both studies also professed to analyzing data fairly frequently. For example, nearly all of the IFL principal survey respondents reported that they examine student achievement data on a weekly basis. Interviewees across both studies similarly reported using data on a regular basis. Several recited common mantras such as, "We are completely data-driven," and "We base all our decisions on data."(Ikemoto & Marsh, 2007, p.106)

The researchers proceed to discuss how it is likely that these statements may mean different things to different people and that "there was not a common understanding among educators of exactly what DDDM entails, or a sufficiently nuanced vocabulary for them to describe various processes and activities in which they were engaged " (p 106) They further demonstrate through case examples how it is likely that communication about DDDM and the development of common understandings is challenged by this lack of consensus.

The researchers provide a number of possible factors that may account for the way data is or is not used based on a synthesis of research (supported by their case analyses). These factors include: the accessibility and timeliness of information, perceived validity of information, staff capacity, including training and the organizational investments in data use, the availability of organizational tools such as regular programs and technologies, organizational culture/leadership, the policy context (including the delegation of substantial implementation details to the states), and (as Coburn, 2004 reported) the fact that external organizations are often part of the systems that provide these data to practitioners.

<u>Barriers to Teachers Using Data (r)</u>

While the study by Ingram, Louis and Schroeder (2004) is titled *Accountability Policies and Teacher Decision Making: Barriers to the Use of Data to Improve Practice,* the barriers to the use of data may exist in the practitioners, limitations in the data, or both. This is a longitudinal study that selected schools from around the country using the

Total Quality Management (TQM)/Continuous Improvement (CI) frame popular in businesses and management techniques. The study begins by stating that the assumption of standards-based accountability is that examination of data will lead to a positive change in teaching and learning. In this study, information is divided into systemic and non-systemic/anecdotal categories. The researchers report that their participants' decisions are as often made through anecdotal and non-systemic information or using a combination of the systemic and non-systemic than are made using systemic information alone. Further, decisions made about teacher effectiveness are reported to be connected to less systematic data, such as course feedback or attendance, than to decisions about school effectiveness. This is further support for the basic proposition that the decisions taken by educators are often complex and that the type of textualized information available from systemic sources is often partially useful for those decisions.

In a number of ways, this study reinforces the notion that issues related to what occurs in classrooms and what occurs in the rest of the educational system require different information and approaches. This view that these dual domains of education – the classroom and the system – require different types of information was indeed shared by teachers as well. The authors note that "This suggests that teachers are not averse to using systematic data for decisions, but that they are more likely to do so when they are making school-wide decisions than individual decisions in their classrooms." (Ibid, p. 1272.) They also speculate that some teachers commonly dismiss externally generated information as a cultural artifact. One of the challenges in this research is to show a connection between a commitment to TQM/CI and the use of data for decisions, further complicating its initial premise.

From the perspective of the teachers, this study details mistrust of data, the differences between what teachers want to know and what is easy to measure, and the general lack of time available to teachers to collect and analyze data. Teachers see their job as preparing their students for life, as opposed to preparing them to perform on various tests. These researchers cite seven barriers to effective use of data in schools. Some of these barriers are cultural; including teacher individual metrics; professionals (teachers and administrators) who have historically based decisions on information that is anecdotal and experiential rather than systematic; the lack of systemic consensus for common measures, and teachers who disassociate themselves from their student performance. Some of these barriers are also technical: what teachers want to measure is rarely available, and schools do not allocate time for the practices that would use data. Finally, they support the historical view that past use of assessment information that leads to a climate of distrust may be a factor as well in the responses participants gave.

Sub-cultures of Beliefs Developed Over Time (s)

The life histories of professionals and the effects of their previous exposure to accountability has appeared in three studies previously reported: Coburn (2004), Falk & Drayton (2004) and Ingram, Louis, and Schroeder (2004). Coburn and Talbert's (2006) study is one that looks broadly at how four organizational levels within one California district perceive what counts as valid research and evidence. By considering both research and evidence, they are connecting student performance information to what has been referred to in other settings as *knowledge* for school improvement (Louis & Dentler, 1988). Their study divides the district into upper and front-line administrators and then combines this with the school, which is divided into principals and teachers. The result is four contiguous organizational levels. This design then allows a consistent lens for the analysis of school and district and the beliefs leading to their title of *Conceptions of Evidence Use in School Districts: Mapping the Terrain* (Coburn & Talbert, 2006).

This study shows that some individuals' beliefs in research and beliefs in evidence do indeed vary in this district by organizational level. The district staff had stronger faith in research than those in schools did. Also, principals in this study notably had a marked skepticism towards research. Further, in their views of the value of evidence, top-level district administrators and principals favored evidence with psychometric properties more than their counterparts. Lower-level administrators and teachers favored assessments that exposed thinking more than the others in the study did. Interestingly, by a wide range, principals gave the highest ranking to multiple measures than any other group did. Further, in considering appropriate use, all groups favored using evidence to support instruction, while only those in schools were in favor of using evidence for placement and only top level administrators were in favor of using evidence for meeting accountability demands.

Perhaps most important from this study is the researchers’ exploration of how similar ideas are shared by individuals in organizational units and how those beliefs can be related to the history of reform messages that the schools have received over time:

> Each of these reform movements penetrated the district to different degrees and in ways that were related to both the formal organizational structure and informal professional networks. We found that individuals who were connected to particular reform movements—because of their organizational division and/or their professional network—tended to hold a similar cluster of conceptions as those promoted by that movement. (Ibid, p. 488.)

The organizational structure was also implicated in this relationship, as organizational units acted like subcultures within the overall organizational context. This extends the general cultural properties discussed previously to sub-cultures related to the bureaucracy’s community of practice. Rather than a culture of inquiry, this work suggests a multi-cultural model may be more appropriate.

<u>Reflection: Topology for Information in Educational Decisions?</u>

After the first two groups of research, the organizational and artifactual, I presented some broad themes that could be found across them. The third and fourth groups allow this conception to be extended somewhat, although the same type of caution I presented earlier about the potential imprecision in these broad categories continues to apply. I will present two types of syntheses we can develop from the problem-focused research and the explanatory frameworks groups of articles. The first synthesis is an elaboration of some of the concepts used in the formulation of the concepts of artifact and organization that these publications allow us to see more clearly. The second involves some ways to think about what these studies tell us about the boundary and interaction between the organizational and artifactual.

Although some of the publications I presented in the first two groups, for example Young (2006), whose study looked at teachers within districts, began to explore the general conceptual nature of educational systems, in these later two groups the different ways that researchers looked at different systemic components can be used to develop a richer model of the systemic construction of educational organizations. In Coburn and

Talbert (2006) four positions that can be thought of as contiguous levels in the educational system are analyzed, and Ingram, Louis, and Schroeder (2004) highlight the specific nature of the role of teachers in the systemic hierarchy. The conceptual nature of the educational enterprise is then divided in another way when authors take up the concept of culture and sub-organizational groups. Coburn and Talbert (2006) let us think of enclaves based on organizational groups with group histories, which they call a diachronic view. A diachronic view is also discussed by Falk and Drayton (2004). Both Diamond and Spillane (2004) and Confrey and Makar (2004) show how accountability status (strongly related to student social class) creates pressures on organizations to dispose them to different types of decision processes. When considering organizations, rather than using general categories, these literatures suggest it is appropriate to refer to their status and their cultural historical construction as well.

In considering conceptually the relationship between the data and the decision – an area that we can call a special case of the more general concepts of symbol and interpretation or text and context – these authors begin to provide some particular discussions of the reciprocal semiotic relationships that can exist between them. While Ikemoto and Marsh (2007) give us a classificatory approach to the decision and the data, other treatments show a directional influence. Confrey and Makar (2004) and Diamond and Spillane (2004), in exploring structuring categorization, show how designations of special education and the subgroups coded into the accountability landscape can have a structuring, if not constitutive, effect on data decision practices. Halvorson (2003) and Massell (2001) discuss ways that these data systems are not just influenced by leaders and superordinate organizations, but that these influencers can direct the creation of these artifacts.

The new understandings not only about information and decisions in education, but about educational organizational structure that these authors are developing are only the surface of what could be an expansive intellectual area. Even if the educational enterprise is some slowly changing monolith with underlying universals, much like the object of geology, as it is often represented in theories of educational organization, this work represents just a beginning of our understanding of it. This literature in both its currency and its nature, however, suggest an aspect of education that is changing.

What Can We Learn From this Literature Base?

What can this literature tell us about how information may be used in Michigan for seventh grade science? In a normal science, the goal of a literature review is to help situate the questions that a specific study should ask, to frame the contribution to the paradigm's literature that the study can make. While in this area there are some general markers that the literature provides, such as that practitioners may tend to over-report their use of information (Coburn, 2004; Ikemoto and Marsh, 2007) and that there are broad topics for critical factors, purposes, and advanced data use, summarized in Table 2.5, there is also significant variation across this literature base. That variation, including the systemic aperture discussed above, make the collection of studies fragmentary and raises questions about whether one study from one systemic area and subject area is comparable to others in other subject areas and focusing on different systemic components. To emphasize the challenges in using this literature base as a substantial frame for empirical research, I will present two additional discussions about the studies reviewed above. The first is what can be called a *potential for topical ambiguity.* The second is *variation in empirical approach. Both* of these support my contention that the research I reviewed, while often using similar terms and developing similar findings, is too immature to be a complete source for the framing of this study.

Topical Ambiguity

When considering the kind of literature that this study explores, the existing literature suggests possibilities for topical ambiguity where publications with similar titles and abstracts may in fact be documenting different types of processes. For example, Streifer's (2002) *Using Data to Make Better Educational Decisions* could either relate to the work that teams of teachers do that Coburn (2005a) and Young (2006) describe, or the decisions taken by school leaders discussed by Diamond and Spillane (2004) and Boudett, City, and Murnane (2005). But, as we know, the contexts that all these authors discuss and the contexts covered by Streifer are often quite different. One reason for this characteristic of some of this literature is that terms such as information or data can refer to texts that may be relevant to many situations. Another is that there is a high degree of reusability with information, so that one textual artifact, for example a student's test score, can be used across different organizational structures and social practices across time, contributing to multiple activities.

Different researchers classify information in quite different ways as well. Some identify information by a measurement approach (ex: running records); others identify information by category (systemic or informal/anecdotal); some by property (ex: psychometric); and still others by its organizational locus (internal/external/school-based). There is additional ambiguity in that certain terms, such as leaders and organization, may often relate to slightly different entities. Is a school leader a principal, a district administrator, or someone who operates as an informal influencer of decisions and others in a school, as Spillane (2006) has discussed?

Variation in Empirical Approach

Variation in empirical approach refers to the differences in research method employed across this literature. While the most common method reported is some type of case study, either formally defined as Young (2006) and Coburn (2001, 2005a) do, or as an example set within a less structured research genre, the pattern of presenting the ways that information is used in educational practice is almost always done through case studies. Table 2.8 reviews the 15 studies (all except the explanatory frameworks) of how data can be used in practice presented earlier with a classification of the nature of the methods disclosed, the type of data set used, and the evidence presented. For the methods column in this table, I have used categories for different levels of transparency. The lowest is a *sketch*, where the authors describe in very general terms the nature of their study approach. The next category is called *described*, where the method is documented in terms of steps and activities but lacking in detail, such as the description of the number of data collection opportunities per participant or details about instruments. The highest category in this column is *explicit*, which is a level of detail usually found in peer-reviewed journal articles.

The second column describes the data set used. A *specific* data set is one that was intended for the literature, while *mixed use* refers to a data set originally intended for another use, but repurposed for the study of data use. A data set that is categorized as *other* is one that is not reported in terms of information use specifically, but that contains documentation of some information use within the data set.

Table 2.8 - Empirical characteristics for studies documenting data use practices

	Study	Methods	Data Set	Evidence Type	Evidence Depth
Organizational	(a) Streifer (2002)	Sketch	Specific	Examples	Deep[2]
	(b) Boudett, City & Murnane (2005)	Sketch[1]	Specific	Examples	Deep[2]
	(c) Supovitz & Klein (2003)	Described	Mixed use	Examples	Medium
	(d) Kerr, et al. (2006)	Described	Mixed use	Descriptive stats/ examples	Light
	(e) Young (2006)	Explicit	Specific	Examples	Deep
Artifactual	(f) Meiles & Foley (2005)	Sketch	Specific	Examples	Light
	(g) Brunner, et al. (2005)	Described	Specific	Descriptive stats/ examples	Medium – deep
	(h) Wayman & Stringfield (2006)	Described	Specific	Examples	Light
	(i) Chen, Heritage, & Lee (2005)	Sketch	Specific	Examples	Light
Problem Focused	(j) Coburn (2005b)	Explicit	Other	Descriptive stats/ examples	Deep
	(k) Coburn (2004)	Explicit	Other	Descriptive stats/ examples	Deep & diachronic
	(l) Diamond and Spillane (2004)	Explicit	Other	Examples	Deep
	(m) Halverson (2003)	Sketch[3]	Other	Examples	Deep[2]
	(n) Falk & Drayton (2004)	Explicit	Mixed use	Descriptive stats/ examples	Deep
	(o) Confrey and Makar(2004)	Described	Secondary	Narrative with examples	Light

1 – More details provided in an article on formative assessment.
2 – Depth inferred but not explicit in literature.
3 – References other larger study where methods are discussed.

The final columns in this table describe the nature of the evidence that the study uses -- both what the authors provide in the literature in terms of evidence and a qualitative judgment about the level of depth of the study. A light study, in this classification, is one based on single data collection opportunities, such as a survey or single site visit. A deep understanding would involve multiple data collection opportunities and multiple data sets, with a medium depth being somewhere in between.

What this brief analysis shows is that, in addition to variation by subject, grade, and organizational role, this literature is uneven in terms of its empirical qualities, with those studies focusing on information use specifically tending to have less transparency and using less robust data sets or data sets that may have been collected for other purposes, while the more standardized empirical reports tend to come from studies that are exploring issues that can include information use, but are focused on other types of challenges.

Adding an Historical Frame: Looking at Citations, Digging for Coherence

If we analyze the literature as it is, then as I discussed above, there are serious questions about its utility to serve a normal science function of framing the study. And this is perhaps to be expected. There were no journals dedicated to these topics nor special interest groups in the American Educational Research Association (AERA) for these issues, although the institutions that many of the authors came from are in the top tier of educational research and the publications and publishers that these references came from are also among the top tier in terms of educational scholarship. Also, as Table 2.5 showed, these diverse articles were often raising similar broad themes. To search for a coherent explanation for the variation of this literature, I analyzed the citations to look for a common core or explanation for the nature of an investigatory community.

The first bibliographic analysis of these publications is to understand to what types of literature they cited. This analysis might help to characterize a movement that would look something like a new disciplinary community. If there is a significant body of common citations across this literature, then it could be fair to infer there are shared perspectives and influences. If the common citations are small, then perhaps what is occurring is a confluence of activity: a mixing of traditions that, while more difficult to characterize as simply artifacts of individual researchers, could be suggestive of a broader movement.

Table 2.9 - Shared citation counts

	Number of publication where a reference is cited					
	1	**2**	**3**	**4**	**5**	**6**
References	437	78	10	2	1	1
Percentage	82.6%	14.7%	1.9%	0.4%	0.2%	0.2%

As Table 2.9 shows, across the fifteen studies, there is not a great deal of commonality of citations, with over 80% of the references used unique to only one publication and less than one percent used in four or more of these sources.

To further understand the nature of these sources, the references were qualitatively coded along two dimensions: the citation type or category and the citation field, which indicates the professional community the reference comes from. The categories relate to the type of resource and included theoretical pieces, research methodology, and research, along with several other categories. Table 2.10 displays the distribution of citations by category, showing that research methods are cited in only some studies, while all cite some theoretical resources. Consistent with the variation in method shown in Table 2.8, those sources that were explicit about their methods were also citing methodological literature, while the others were not. In general, this group of studies, however, is theory rich and methods poor.

Table 2.10 - Number of theory and research citations by source.

	- - - Sources - - -														
	Organizational					*Artifactual*				*Problem Oriented*					
	a	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i**	**j**	**k**	**l**	**m**	**n**	**o**
Research	6	34	24	5	17	3	21	34	25	35	37	35	41	34	10
Res. methods	0	0	0	0	4	0	0	0	0	7	4	5	1	1	0
Theory	8	9	3	4	10	6	11	9	4	19	19	9	8	27	3
Other*	7	3	5	7	2	14	10	3	1	10	14	15	7	2	9

* includes speeches, references to source material, and to previous work in research program

This qualitative coding of the citations further supports a view that the organizational and artifactual investigations are very similar to each other and qualitatively different from the problem-focused literature. With the exception of Young (2006), none of the sources in the first two categories have explicit reference to research methods, while all of those in the problem-oriented category do. Young's work, while topically more consistent with the first group, was conducted in a tradition similar tradition to Coburn's. While two of the problem-oriented sources have only one methodological citation each, these two come from the Distributed Leadership Project (Spillane, 2006), which involves an important methodological component developed to understand leadership practices in schools. Possibly several of the Distributed Leadership references in these works acknowledge methodological choices. Across the information use literature, there are some cases where the reference to previous research is light, where all of the references in the problem-focused literature in this review includes significant reference to prior empirical work.

Another citation analysis looked at the fields that were influencing the researchers according to where their references were obtained. The percentages of the overall citations in each publication, coded by professional field, are presented in Table 2.11. These results indicate some of the different traditions being brought together in these studies. Within all three of the categories I used in grouping the research articles – organizational, artifactual, and problem-focused –there is substantial referencing of teaching and learning, educational leadership, educational organization/administration studies, and policy and accountability studies. This cross-section of disciplines suggests some of the systemic issues that this group of literature addresses as they touch on many important parts of the educational enterprise.

Table 2.11 - Percent of citations coded for disciplinary orientation

	Organizational					***Artifactual***				***Problem Focused***						
Topical (field) code	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	All
Policy & accountability	-	8	3	5	6	-	31	8	6	23	32	3-	14	7	13	16%
Education org./admin	-	15	12	10	17	8	2	15	15	11	11	10	9	28	4	12%
Teaching and learning	-	19	6	2-	6	-	2	19	6	12	13	15	23	10	-	12%
Data use (organizational)	14	13	58	-	17	21	16	13	36	-	-	-	-	-	26	10%
Measurement/assessment	5	26	-	15	6	4	22	26	3	-	-	11	2	-	13	8%
The teaching profession	-	6	-	10	9	-	2	6	-	12	3	10	-	14	-	6%
Organizational Science	5	-	-	10	6	-	4	-	-	12	16	-	5	10	-	6%
Data use (artifactual)	5	4	15	-	9	46	11	4	21	-	-	-	-	1	4	5%
Social issues and theory	-	-	3	5	-	4	-	-	-	5	1	-	24	7	-	4%
Research methodology	-	-	-	-	11	-	-	-	-	8	5	7	2	1	-	3%
School leadership	-	2	3	-	6	-	-	2	9	-	-	-	5	11	-	3%
Cognitive science/psych.	5	2	-	5	3	-	-	2	-	1	3	-	2	6	4	2%
Information science	48	-	-	-	3	4	2	-	-	-	-	-	-	1	-	2%
Educational technology	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	<1%
Other	19	6	-	2-	3	13	4	6	3	16	16	18	17	4	35	11%

The similarities between the organizational and artifactual orientations discussed above in Table 2.5 are reinforced in this analysis, as both types of sources tend to draw upon both types of publications. Further, perhaps as a result of the general rather than technical nature of the publication venues, information science and educational technology are not widely prominent, despite the fact that technological infrastructure and capacity was reported as a critical factor for success (see Table 2.5). Further, there is no representation across this body of work for symbolic communication, despite the indications that visual presentation of information is critically important to success as well.

The most frequently-cited category of research shown in Table 2.11 is policy and accountability literature. This reinforces the strong emphasis on NCLB across these studies. Fourth on the list is the field of information or data use from an organizational perspective. This is the first group of studies reviewed in the section on Recent Empirical Work. Further down the list is data use from an artifactual perspective, which was the next group of studies I reviewed in detail. These two combined would create a category that would account for 15% of the citations in the literature, or the second highest category. This indicates that while these authors are largely drawing from different sources and from a broad range of disciplinary traditions, the sources they do share may be coming from within the field itself.

To understand this possibility, I looked at cross-citations, where literature that I reviewed was citing other literature I reviewed. Given that the body of literature is so new, it would also be important to understand how there may be an accumulation of cross-citations and whether some of the publications are beginning to take on characteristics of foundation literature that would be often cited as part of an emerging paradigm of inquiry. Figure 2.1 illustrates this principle by showing the different types of cross-publication citations found within the literature reviewed earlier. The interconnections between these sources may be a sample of the interconnections that exist in a more comprehensive selection of literature. They show that even though each of these sources is largely drawing off of different literatures (see Table 2.9), there is evidence of an emergent scholarly dialogue. This dialogue primarily involves the data use research, while the problem-focused research plays a contributory role.

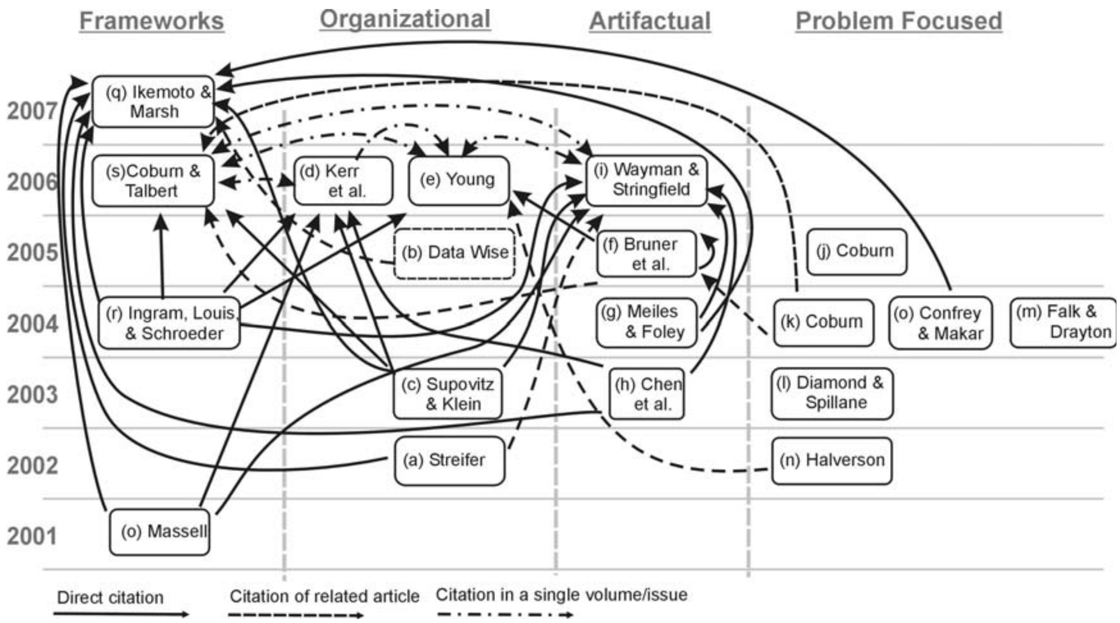


Figure 2.1 - Cross-publication citations

In addressing the larger question of how this emerging literature fits into an historical progression, one possibility is that these investigations are the beginnings of a foundation for what will become a professional community pursuing coherent programs of inquiry. The accumulation of cross-citations over time may indicate this possibility. Within this hypothetical view, might some of the more comprehensive, agenda-setting, data use literature be used to frame more specific problem investigations than the patterns we see today? If so, then this new field would be one of the most interdisciplinary combinations seen in education, with representation from leadership and administration, policy, teaching, information visual design, and technology, to name a few. It would be a field of inquiry with direct connection to important elements of the standards-based reform movement and one that, with the advance of technology and information use across so many professional contexts of educational practice, would be likely to have continued opportunities for productive research.

This field would also be one that is facing some methodological challenges. Perhaps studies into the use of data present some special research complications because, unlike instruction that occurs generally in one location (the classroom) and is mostly continuous day after day, data use is episodic and crosses contexts. Methodological challenges may be inherent in this domain of study, and the literature reviewed here could represent initial repurposing of research methods that have been proven useful in studying other types of issues, but perhaps are not the ideal methods for systemic study.

Building on the Emerging Literature: a Paradigmatic Perspective

Earlier, I presented three goals for this literature review. The first was to describe the empirical nature of these activities. The second was to historically characterize the domain of investigation. And the third was to identify how this study could build upon this literature, both in terms of theoretical tools and the empirical contributions the study can make. The literature I reviewed has a curious characteristic. It often covers different parts of the educational system, but is beginning to show similar types of broad themes. It draws widely from what could be considered different types of sub-fields in educational research, but has very little overlap of source literature. However, the overlap that does exist includes active cross-referencing that can indicate the development of a research community and/or an intellectual *founder's effect*,[8] where a small number of ideas from early works gain currency across the literature. Since these studies, while addressing similar topics, do not yet have the qualities of a normal science, they can be interpreted as the beginnings of a research paradigm that Kuhn (1970) described as the historical framework for scientific communities.

Taking a paradigmatic perspective, the contributions that this study can make can be seen not only in the area it studies – middle school science in a Midwest state – which are all underrepresented in the literature, but also in its approach and the ideas it brings to the discussion. One reason for the variation in this literature is that it is exploring issues that are by their nature *systemic*: they span contexts and temporal periods. As a result, it makes sense that there would be such variation in systemic aperture, because researchers would select segments of the system that have a history of being studied -- for example, a district or a school -- and locate within that segment of the system findings that relate to other researchers who are investigating other aspects of the system. One perspective that would add to this literature is a systemic theoretical perspective that can integrate

[8] A founder's effect occurs in populations of ecosystems when a small number of organisms of a species populate a new environment (ex: a volcanic island) and the gene from those original founders then becomes the basis for a gene pool with low levels of diversity.

different contexts and practice dimensions. Another contribution of this study is a body of research that helps to advance the genre of this field by combining the broad agendas of the organizational and artifactual perspectives with the methodological rigor of the problem focused literature.

Chapter 3 Some Theoretical Tools

> When you do buy into a theory, Kuhn continues, you ‘begin to speak the language like a native. No process quite like choice has occurred’, but you end up speaking the language like a native nonetheless. ~ Ian Hacking

As Hacking (1983) reminds us, the theoretical fabric of a scientific paradigm can become a language that its community learns, uses, and perhaps makes incremental modifications to without the formidable task of designing, constructing, and assembling it anew. For mature and established fields, of which in education there are many, the theoretical roots drawn upon are often set deep. This study addresses a new area that Moss and Piety (2007) describe as having boundaries that are “very much under construction” (p.1). I will tentatively name this area *educational systemics* because of its interest in how different individuals from different types of educational organizations work together. This collaborative work is often constituted by and features information (data) about student performance related to national policy. Data or evidence, rather than being an object in itself, can be a mechanism for common work that crosses boundaries around the goal of improving student performance.

> When evidence crosses boundaries, it brings far more than information: it entails sets of cultural tools—including artifacts, concepts, and often norms and routines--that mediate understanding and (inter)action in sending and receiving contexts. In the case of large scale standardized assessments, for instance, the set of cultural tools includes artifacts like stated goals of the assessment, test forms, standards or domain descriptions, guidelines for evaluating performances, score reports, technical manuals, regulations for users, and so on; it includes concepts that represent what is important to learn and what counts as evidence of that learning; it includes expected practices (rules and norms) like standardized administration formats, independent work, and so on; and it entails an implied division of labor (different roles for test developers, teachers, students, parents, administrators, policy makers, and others) in the construction and use of the information. (Moss, Girard, and Haniford, 2006, p.145)

The previous chapter showed a field that was new and drawing on many different traditions (see Table 2.11). The goal of this chapter is to provide a set of theoretical resources that have been useful in the design of this study and interpretation of its findings. As this study attempts one of the widest views into this new field -- one that addresses processes at the level of a state -- I will out of necessity use a theoretical pidgin[9], a small set of tools, that I can carry in what is essentially a cartographic quest. An important element of this study is a rough mapping of systemic structure in Michigan relevant to my central topic of seventh-grade science educational assessment. This mapping is a key component in a description of the systemic interactions around the use of student performance data.

The conditions under which this study was conducted played a large role in using a reductionist and descriptive approach called Actor-Network Theory, or ANT (Latour, 2005; Law, 1992), as a principle tool. ANT has several important properties that make it suitable for this effort. First, it looks at actor relations and associations in a particular way. This is an important element in the question this study addresses about the role information plays in collective or group activity. The collectives and the groups need to be understood in order to understand the role the information played in their associations. Further, ANT is a model that does not prioritize hierarchical relations and would not be used to create an *a priori* nested structure. Rather, it requires that encompassing relationships be empirically demonstrated. The study context that I provide below will illustrate that Michigan's educational enterprise does not easily fit into a purely hierarchical model as it contains multiple overlapping relationships between the schools and the state. ANT also puts objects or artifacts in an important role. It considers their effect on practice to be similar to the effects that humans can have. This is important as this study includes attention to assessment artifacts as well as practices.

In this chapter, I will first present some classical and recent views on the nature of educational organizations. I then have a section that describes the organizational topology in the State of Michigan. This may help readers to orient themselves to some

[9] A pidgin is a simplified language that forms across two different language communities so that thy can communicate. My use of it in this context is rhetorical.

aspects of this study not represented in much of the literature reviewed in Chapter 2. That section also discusses some important aspects of Michigan's standards and assessment efforts relevant to science education at the time this study was conducted. Following this discussion of the State of Michigan, I discuss ANT in more detail so as to provide a short synopsis of the characteristic features that influence its use here. I will then present some other theories that I draw upon and provide definitions for key terms I will use in the dissertation.

Some Foundational Views of Educational Organizations

The organizational nature of educational systems continues to be a matter of active concern for educational theorists and researchers. How is the educational system in the United States similar to or different from the U.S. Army, the Catholic Church, a supermarket chain, or the health care system? This is a fundamental question for policies and innovations that seek to reform or influence the behavior of education systems. Of course, to ask this question presupposes the existence of a system, a proposition brought into question by Cohen (1995) in his classic *Where is the System in Systemic Reform?* Cohen discusses the myriad of entities, both public and private, that are involved in matters related to schooling. The conceptual difficulties that Cohen highlights, however, have not prevented many before and since from proposing conceptual frameworks to describe the fundamental character of the educational enterprise.

Traditionally, educational organizations have been considered in a hierarchical arrangement (Green, Ericson & Seidman, 1980). And there are certainly elements of the educational system that are hierarchically arranged. Students within classrooms and classrooms within schools are ubiquitous structures, although services associated with special education populations may complicate these structures. Further, recent research has supported hierarchical views by providing evidence that school level differences can significantly account for differences in student learning (Lee & Bryk, 1989; Raudenbush & Willms, 1995). This challenged a previously supported position that students' social environments were overwhelmingly responsible for schooling outcomes (see Coleman, 1966) and perhaps led to renewed interest in understanding the differences between these hierarchies.

In an early and still often-cited work, Weick (1976) describes educational organizations as *loosely coupled systems*. This model theorizes poor coordination between roles (ex: principals and teachers), the preservation of individual and sub-group identities (ex: math teachers), and resistance to change. This model helps explain why many policy initiatives ultimately fail to produce appreciable effects on instructional practice (the *technical core*) of schools. The metaphor of loose coupling has had a strong effect on the study of school organization. Thirty years ago, Weick said:

> At the outset the two most commonly discussed coupling mechanisms are the technical core of the organization and authority of office….A compelling argument can be made that *neither* of these two coupling mechanisms is prominent in educational systems found in the United States. (Emphasis added, p.4.)

In other words, what Weick argued was that neither the central authority of schools (the chain of command from district to school leaders) nor the core activity, classroom teaching, was sufficient to bind the educational organization together. Exactly what organizational coupling is and how it can be studied is a matter of discussion. Orton and Weick (1990), for example, identified four competing approaches to coupling theory: counter-rational, unidimensional, multidimensional, and dialectical. The last approaches account for multiple linkage mechanisms with variable strengths. Weick (1982) further defined coupling in terms of the direct and predicable nature of interactions. Young (2006) applied a loose coupling model to both the organizational action (practice) and to the documents (artifacts) that are part of that action.

Another view of educational structures illustrates that educational systems are not always representative of functional responses to what the organization requires. Schools can be seen as institutions that reflect the desires of their stakeholders, as they need to achieve legitimacy, and that also are shaped by efficiency demands (Meyer & Rowan, 1977). Also, when considering organizational aspects such as school leadership, there can be significant differences in the formal designed leadership systems and those that are lived within the school (Spillane, Camburn, Lewis, & Stitziel-Pareja, 2006). Further, while the organizational structure seems intuitively related to instructional practices, research in some cases has shown weak relationships between the two (Elmore, Peterson, & McCarthey, 1996).

Between the time that Weick first wrote about the educational system and today, some developments have taken place in American education that are important for this research. In science education, national standards have been adopted that have influenced local standards (AAAS, 1993; NRC, 1996). Also in recent years, the science education field has seen institution-building in the area of science specialists (see http://ecommerce.nsta.org/bap/). More generally, there have been efforts to redefine the role of principals as instructional leaders rather than as building administrators (CCSSO, 1996; Fink & Resnick, 2001). All of these aspects of recent conditions can be seen as developments that are designed to lead to greater cohesion in the educational system and to broaden instructional responsibility beyond classroom teachers.

Subject matter may also influence differences in schooling practices. In revisiting the theory of loose coupling, Burch & Spillane (2005) argue that despite the explanatory power of the loose coupling model, schools are not universally loosely coupled, and the strength of relationships is often related to the *domain of instruction*. When instruction is treated as a unitary phenomenon, important differences between different subjects are not recognized. Teaching means teaching *something*, so it follows that the area of assessment and instructionally-related decisions in the policy-to-practice relations of mathematics might differ from, say, science and literacy.

The location and sector of the school also matter. In science education, variation occurs between schools located in different areas. Across the various states, there are both differential levels of commitment to standards-based reform generally (Carnoy & Loeb, 2003) and great variation in the adoption of science standards (Lerner, 1998). Research has also shown that there are characteristic differences between the instruction that occurs in public schools versus parochial (Marsh, 1991), in large versus small (Lee & Loeb, 2000), and in urban versus rural (Martin & Yin, 1999), to name a few of the axes upon which the technical core of education can be analyzed.

Recently the field of educational research has also seen studies that are exploring the network properties of educational organizations. Spillane's Distributed Leadership Study (2006) explored network leadership structures that parallel the official hierarchical structures. And, recently, social network analysis (Scott, 2000) has been used by Franks and Penuel to map professional relationships with the aid of advanced quantitative tools (Frank, Zhao, & Borman, 2004; Penuel, Sussex, & Hoadley, 2006). However, much of this work has focused on teacher communities and schools, rather than across organizational boundaries in a given state in the way that this study does.

Particularizing ANT for an Educational Study

Actor Network Theory, some might say, is really not much of a theory. While it is the principle guiding framework for this work, it is not a theory in the predictive sense. It is not the kind of theory that one can operationalize and empirically validate and/or elaborate upon. It is, rather, a research framework that seeks to avoid several problems that have plagued sociological studies (Latour, 2005). Growing out of a tradition of social studies of scientific practices with roots in anthropology (see Latour and Woolgar, 1979; Latour, 1987), the originators of ANT began to look critically at practices of sociology and the ways that certain analytic categories such as *actor* or *society* were used by researchers. While ANT does involve network structures and relations, its primary thrust, according to Latour (2005), is to provide a way to approach these analytic categories empirically and particularly.

The foundational move in ANT is similar to some sociolinguistic approaches. Actor and society can serve in roles similar to the role that *context* has had in some communication theories: it is a constraint on language and semiosis. Also coming from anthropological traditions, scholars in the emerging sociolinguistics field challenged the use of context as an enclosing, constraining, catch-all for meaning-making (Goodwin and Duranti, 1992). In Latour's (2005) ANT primer, *Reassembling the Social: An Introduction to Actor-Network-Theory*, more than the first half of the book is devoted to what an analyst should not do. The second part of the book is devoted to the painstakingly detailed work of tracing associations between actual real world entities, rather than relying upon the convenience of black box categories such as actor, society, and even system can be. One can think of ANT in some respects as an ethnography that travels. Where ethnographies are usually deep studies of a cultural locus, ANT studies can traverse network relations while retaining an intimate investigatory approach.

In this short overview I will focus on three aspects of ANT as it is being applied to this dissertation. The first is its approach to the study of power relations, which could be implicitly associated with issues of scale (e.g., the state is both larger in scale and exerts broad power over lower subordinate levels). The second is its stance on universality, typicality, and uniqueness. The third is the approach in ANT to dealing with objects or artifacts. All three of these are relevant to this study, as educational research generally and the study of assessments in particular have often featured strong attention to typicality. Whether classroom quiz or the state test, I am studying social relations where these objects (texts) may play instrumental roles.

To address issues of power and scale, Latour (2005) recommends flattening the network by using the same lens to analyze across the different loci of the research. "Size and zoom should not be confused with connectedness" (p.187), he advises. He stresses that in tracing a network of relations across different types of sites, differences do exist. However, those differences, including those with power relations, need to be empirically demonstrated rather than assumed. He says that "…framing activity, this very act of contexualizing, that it should be brought into the foreground and that it cannot be done as long as the zoom effect is taken for granted" (p.186). This principle has important consequences for this study. When including the state test within the same view as more localized entities such as classrooms the division or linkage should be documented, and this documentation should extend into the locus of the government operation and not stop outside its door. I studied the state's assessment development process using the same methods I used in schools. We also would not automatically place teachers within a nested structure, but rather would be required to empirically show how their practice is enclosed or constrained by their schools.

In looking at individual actors, Latour includes both the human and non-human in the same general category. While not going as far as giving objects and technology the same qualities as human beings, ANT does consider that in many situations these non-human *actants* can have the same types of effects on situations that humans do. Rather than suggesting technological determinism -- the belief that technologies actually cause certain outcomes -- Latour (1996) refers to the approach as "a semiotic definition -an actant-, that is, something that acts or to which *activity is granted by others*" (emphasis added, p. 5). For this study, the implications should be clear: the artifacts that are used in assessment need to be included in the study and their material relations to other actors in the study considered as well. We look at them not as static texts, but as being enmeshed in complex social practices (Swales, 1998). According to Latour (1996) "…actors are not conceived as fixed entities but as flows, as circulating objects, undergoing trials, and their stability, continuity, isotopies has to be obtained by other actions and other trials" (p. 6).

The third issue I want to highlight is the issue of typicality or uniqueness. The starting position of ANT is that commonality needs to be explained and not assumed. This is consistent with the rest of ANT, but it is important to emphasize, especially given that so many of the studies reviewed in Chapter 2 discussed what various types of actors in different positions believed about data or used data for. Just as different red blood cells and different red balloons may take different paths and yet follow in form and function their prototypes, it is convenient to consider different types of actors (ex: teachers) or different kinds of artifacts (ex: a classroom test) as part of general categories that can be used as building blocks for larger assemblages that are then representative of universal truths. ANT clearly rejects the use of types and typicality that have not been demonstrated:

> Universality or order are not the rule but the exceptions that have to be accounted for. Loci, contingencies or clusters are more like archipelagos on a sea than like lakes dotting a solid land. Less metaphorically, whereas universalists have to fill in the whole surface either with order or with contingencies, AT (ANT) do not attempt to fill in what is in between local pocket of orders or in between the filaments relating these contingencies. (Latour, 1996, parentheses added, p. 3.)

For this study, the implications are significant. In tracing individual associations into a network that reaches into the state government apparatus, I am not developing a generalized or typical conception of either its parts or the whole. Rather, I attempt to preserve individual uniqueness in the study and leverage the unique positions that some participants have to expand the study's understanding. I seek to develop more holistic depictions of key events than might be possible with participants that did not have these unique characteristics.

This study departs from the approach Latour has used in his empirical work (see Latour & Woolgar, 1979; Latour, 1987; Latour, 1996) in terms of the framing of the study and the selection of objects of analysis. While Latour will follow actors and the artifacts he calls actants, when he does so he has begun from a single meaning-making endeavor, such as a scientific laboratory or a new transportation technology. In this study, the focus is on a segment of the educational enterprise that is defined by policy. If Latour were undertaking this study, he might have focused on the state government's test development process and followed some of the same members of that bureaucracy through their work in internal meetings and with various stakeholders in the state, as well as their professional community interactions. He might look at how they mediate between federal and local interests and pay particular attention to the artifacts that this group develops in terms of the reports of assessment results (some of which I show in Chapters 8 and 9) and how these *immutable mobiles* (Latour, 1990) circulate through schools and homes.[10] This study, however, uses the state boundary as a perimeter of the investigation. It looks at part of the state's meaning-making process and aspects of educational practice starting in classrooms related to the state test. The principles of ANT are used where a particular and ethnographic approach is applied across contexts. But rather than a single meaning-making endeavor being used to organize the study, a category of meaning-making work – seventh grade science assessments of learning – is used to organize the investigation.

Timescales and Temporal Modeling

Embedded within ANT is a theoretical interchange across scale. ANT researchers not only apply the same research methods to different scalar entities in their own research, the methods are of the same kind as are used in very localized studies of interaction. ANT is about meaning-making activities (Latour, 2005), and if we look at very particular research into verbal, visual, and gestural communication, we find the very same types of questions being asked. In Erickson and Schulz' (1987) *When is a context? Some issues and methods in the analysis of social competence*, the researchers discuss not

[10] This would be a wonderful study.

only how social situations constrain and index meaning, but also how the communication reciprocally shapes the context. Looking at ANT at a social level and Erickson and Shulz's similar approach at the conversational level – both of which deconstruct the semiotic container – we can see compatible approaches across different social scales. This theoretical transport leads to some of Lemke's work on ecosocial levels and timescales.

Lemke's (2000) *Across the Scales of Time: Artifacts, Activities, and Meanings in Ecosocial Systems* makes a number of important contributions to this study. First, it takes the concepts of meaning-making that Latour bifurcated into local and social and generalizes them across multiple scales. Going from very small entities (ex: cells) to very large ones (ex: galaxies), Lemke looks at alternations of topologies (ex: scores) and types (ex: performance groups) across levels. He argues for a study of multiple levels in a system where any level being studied should also entail the study of its immediate higher and lower levels in the ecosocial hierarchy (N, N+1, N-1). This would mean that a study of classrooms should also entail a study of both schools and individual students. While the hierarchical application in the context of this dissertation is complicated, with intermediate levels that are often irregular and inconsistent, the approach is still important, as hierarchical elements do exist within educational systemic collections; and some information constructs, such as the interim assessments reviewed in Chapter 2, can be seen as exemplars of hierarchies across levels in these educational contexts. This theory also posits that one reason why policies or semiotic efforts from one ecosocial level (ex: state) have weak effects on another (ex: a classroom) is that they operate on different timescales.

The concept of hierarchical relationships between levels in educational systems and between different forms of assessment can also be found in more recent literature. For example, Erickson (2007) makes the claim that a fundamental difference between formative and evaluation (summative) assessments is the timescale on which they operate:

> One of the basic problems in relating educational evaluation and educational practice is that the two activities often take place on radically differing *time scales*. It is not only a matter of aims—that evaluation of

> local educational practice as conducted by external researchers (or by the use of instruments designed by external researchers, as in the case of formal testing) may be done "summatively" for purposes of external accountability, and so the information collected may not directly inform the local conduct of instruction and school administration (emphasis added, p. 186).

This approach to looking at formative or classroom practice and accountability practices within a uniform framework is also found in national policy texts such as the National Research Council's *Knowing What Students Know*, which states that "The committee recognizes that all assessment is in a sense "**formative**" in that it is intended to provide feedback to the system to inform next steps for learning" (emphasis in original, p. 38).

Lemke's original ecosocial work has led also to an elaborated model more focused on educational systems represented by Lemke and Sabelli (2008) that, while still retaining hierarchical theoretical orientations, calls for detailed study of subsystem components and their mechanisms for interaction. Even as this study begins by problematizing nested frameworks, it can be seen as falling directly within the research agenda that Lemke and Sabelli recommend:

> Most important perhaps is a change in the paradigms of our thinking about research on education. Away from input-output "blackbox" causal models to modeling the specific, local linkages that actually interconnect actors, practices, and events across multiple levels of organization. (p 128).

Boundary Practices and Related Terms Used in the Dissertation

While this study is largely particular and requires a translation to be applied to broader contexts, the research discourses contemporary to it operationalize concepts that can be useful in that translation process. To help this process, I use terms and rely upon additional theories that include *activity systems* (Engestrom, 1987; Cole, 1996); *communities of practice* (Lave and Wenger, l991; Wenger, 1998); *organizational coupling* (Weick, 1976; Spillane & Burch, 2004); *brokers* (Wenger, 1998); *boundary spanners* (Tushman and Scanlan, 1981; Honig, 2006), and *boundary objects* (Star & Griesemer, 1989; Wenger, 1998). All are different ways of framing a similar type of research question: How do individuals within and across groups routinely connect and interact? Below, I will provide short overviews of these concepts as they relate to this study.

Activity Systems and Communities of Practice

Both activity systems and communities of practice grow out of the sociocultural revolution in psychology that drew heavily upon the work of early twentieth century Russian psychologists, specifically the ideas of Vygotsky (see 1985, 1997). These theories focused attention on groups of individuals in historically persistent and culturally bound groups. Communities of practice theories (Lave and Wenger, 1991) have focused on the process of gaining entry through participation trajectories from novice to full membership. Communities of practice researchers have often worked in out-of-school contexts or with schooling/apprenticeship programs that are dissimilar from public education. Activity system theorists (Engestrom, 1987; Cole, 1996) have focused more on structured settings like classrooms and organizations. While useful and used in some ways throughout this dissertation, these concepts have the potential to become reified, and then could be used in ways similar to concepts like *society* or *actor* to support general notions with often only particularized evidence.

While both of these concepts are important in the field of educational research and will be featured in this study in important ways, from an ANT or Latourian perspective they raise an important question: When these concepts are used, are they essentially middle-scale members of the family of black-box entities to which societies and actors belong in traditional sociological analysis? For this reason, in the design of the study, I will follow ANT by selecting specific organizations and individuals and apply these concepts of activity system and community of practice to particular examples.

Brokers and Boundary Spanners

Brokers and boundary spanners are individuals who cross between groups that can include formal chartered organizations, like companies or governmental units, or less formal structures where insider-outsider differences can be found. Wenger (1998) focused on brokers as individuals who can “introduce elements of one practice into another” (p.105). For Wenger, boundaries and brokers are designed; they can be put into organizational structures for certain purposes. Similarly, Honig (2006) highlighted the role of these individuals who can cross organizational settings in education:

> Public management research in sectors outside education suggests that the designation of central office administrators as boundary spanners holds promise for helping central offices reshape their relationships with schools and other youth-serving institutions. (p. 158)

Honig's research, however, showed that what seemed to be a straightforward linkage between organizations based on organizational role varied by individual and over time. Substituting ANT's fluid and particular approach uses a circulating, dynamic, and particular ANT actor whose relationships with particular organizations and particular individuals are traced over time. These terms, *boundary spanner* and *broker*, will be used in this study to broaden claims somewhat from individuals to groups of individuals, while retaining a focus on particularity.

Boundary Objects

Boundary objects are artifacts whose material properties help to structure activities in the same ways that ANT looks at actants. Star and Griesemer introduced boundary objects to the theoretical lexicon in a study of an early 20th century museum. Wenger (1998) further elaborated this concept, placing boundary objects within a more contemporary context that included information systems, but without the same type of empirical support. Latour's immutable mobiles serve similar purposes through their ability to travel across contexts. There are some important distinctions between the original use of boundary objects as articulated by Star and Griesemer (1989) and Wenger (1998) and the focus of this study. These differences are related to the differences between this study and the types of original ANT studies undertaken by Latour.

Star and Griesemer studied the work of an influential administrator in the natural history museum at the University of California Berkeley who was interested in documenting flora and fauna in California and needed the support of a variety of types of armature and professional collectors to accomplish his goal. To meet the need to view the collection process in multiple ways, including the detailed cataloging and recordkeeping associated with the development of a professional collection, he developed a series of forms and indexing approaches that allowed different actors to participate collaboratively in the process. Central to this administrator's work were the concepts of collaboration and negotiation. The material artifacts he designed were records of that process that included reconciling multiple interests. Wenger (1998) takes a similar approach in the discussion where he is theorizing about the design of information systems across similar types of organizational communities with multiple interests. He cites four essential characteristics of Star's model for boundary objects:

- Modularity, the ability of different portions of a boundary object to relate to different perspectives;
- Abstraction, the object abstracts elements essential for all perspectives and deletes or masks those that are not;
- Accommodation, the object is useful in a variety of activities;
- Standardization, the information in the object conforms to pre-defined meta-rules.

This study departs from these more classic works in two important ways. As Chapter 8 will discuss, the individual state test item, the MEAP item, can be seen to hold a structuring role in discussions that occur in both the state test development process and the local review of test results that schools conduct. It is implicit in what I am referring to as boundary practices (Wenger, 1998) that occur similarly across both of these contexts. However, missing from this process is the concept of negotiation and collaboration. The federal policy has delegated to the state government the responsibility for the design and administration of the accountability assessment process. While Michigan does include practitioners, these practitioners' interests do not seem to have been considered in the design of the assessment artifacts in the same way that Star and Griesemer (1989) and Wenger (1998) envisioned for boundary objects.

The second way that this study departs from the original conceptions of boundary objects is in the role of alternative representational technologies present in this study. In Chapter 8, I show how the MEAP item plays a structuring role in discussions about a range of systemic topics. In Chapter 9, I show how the MEAP item plays a different role, because a different representational technology presents assessment results by different groupings, called *strands*. In these other conversations, the MEAP item has not changed, nor have the formats in which the state presented it. Rather, the social practice changed, in part because it was based on a new way of representing the results, presenting them according to strand. The whole issue of materiality can then be viewed in different ways. No longer can the item be viewed as a material object, even though it may be presented on a printed page. It must be considered a type of text that can appear in different textual spaces and different textual configurations. These are important distinctions for which I have little theory to present.

Learning to Speak the Language

As useful as these concepts are, they are of much greater value when they have been demonstrated to exist in particular contexts rather than as general analytical categories. Each one -- the grouping, the boundary crossing individuals, and the artifacts that link action -- are also part of ANT, but often with different names. For example, brokers (and boundary objects) in ANT could be considered as actants that cross loci and participate in meaning-making. Latour describes these in two categories, intermediaries and mediators; the latter changes meaning, while the former does not. The circumstances under which various individuals or artifacts affect meaning cannot be assumed to be universal, but rather are subject to particular contingencies.

As Hacking's quote at the beginning of this chapter reminds us, the use of theory can be automatic and, like language, once we acquire it we can forget our cultural appropriation. Languages are inseparable from their social practices of use. And this study is related to what, at the time it was conducted. was a set of new and evolving research discourses. This study, by taking a fully descriptive stance, by addressing issues that are related to recent technology and a nascent literature, and by focusing on a part of the United States that is complex and irregular in terms of organizational structure, holds the potential to sketch some features important to this evolving community's language.

Chapter 4 Study Overview: Methodology, Context, Sample, and Analysis Plan

Everything should be made as simple as possible, but not simpler.
~ *Albert Einstein*

Introduction

This chapter describes the study's design and some key analytic concepts used throughout the remainder of the document. It describes in detail the methodology used in this work, called a *network model case study*. As this is a new method and being used to investigate processes not strongly represented in research literature, my presentation will involve some detail about the nature of the method. Following a short overview of the method, I present a more detailed definition of the methodology, including some descriptions of prior work that I believe supports my development and formalization of the research approach. In this definition, I will specify the criteria I used to define various units of analysis, such as cases, and some of the theoretical reasons for those decisions.

I include a discussion of the research context in the state of Michigan that describes its organizational structure and some changes that were occurring across the state during the time of the study. I return again to Michigan in discussing the history of the study: how the study progressed from one that focused on schools across states to one that focused on schools in a systemic and diachronic frame within a single state. This discussion explains the emergent nature of the research (Patton, 2001) and attempts to clarify sampling choices that are artifacts of the research process. An overview of the sample will describe in broad terms the nature of the evidence that will be presented in more detail in the analysis chapters that follow. One characteristic of this study design is that it does have many different types of data and several analysis strategies, owing in large part to the span of organizational activities it is attempting to address (referred to in Chapter 2 as the *systemic aperture*). Because different types of data are used differently within different parts of the study (in different types of cases), I discuss the various categories of evidence and their sources along with their relationships to different types of cases.

The latter part of this chapter includes a section detailing the case-based analysis plan. This plan describes the chain of evidence (Yin, 2002), which begins with coding of artifacts and association of data cases, to show how the analysis approach leverages the network structure of this evidence to target the key questions this dissertation addresses. I have included an overview of the coding structure that is being used in the analysis of much of the participant data. Finally, this chapter concludes with a brief overview of the technology used to support the analysis.

Overview of the Method

The methodology used in this study is based on a methodological category with many variations; it also shares many of the important definitional and analytic elements of more established case study methods. This study's methodological variation and the analysis techniques are based largely on the work of Yin, who states that the case study "investigates a phenomenon within its real-life context" and is useful when the "boundaries between phenomena and context are not clearly evident" (Yin, 2002, p. 13). The extensions I have made to the case study method are necessary to accommodate the nature of this study's systemic aperture, with its crossing of different organizational levels in a system that has both hierarchical and non-hierarchical (network) properties. These extensions are consistent with the general character of the case study as a naturalistic investigative approach with applicability to both description (Patton, 2001) and theory building (Eisenhardt, 1989). As with other case study approaches, this research uses multiple forms of evidence and examines processes that occur over time. It could be seen in a general sense to be a case study design with a foundation based on Yin's framework and substantive methodological and theoretical contributions from Actor Network Theory (ANT).

Case studies are organized around decision processes[11] (Yin, 2002). They have a strong narrative element that allows them to be understood broadly (Stake, 1978). And they have often been applied to educational circumstances (Merriam, 1998). Case studies are built around phenomena documented as discrete cases that can be arranged in many different ways. Some studies have a single case, some compare multiple cases, and some have multiple cases in a nested structure. The network model I am defining for this study adds a dimension to the sub-case relationships. Rather than sub-cases being of the same type and discrete, in what I referred to in Chapter 2 as a cellular case structure, sub-cases in this model can be of different types relating to different sociological phenomena and interrelated in ways that are more reflective of the field of education.

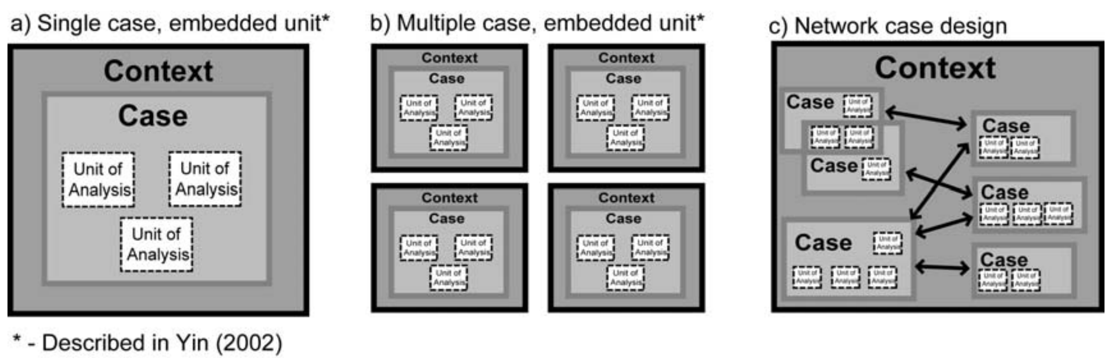


Figure 4.1 - Examples of different case study design structures based on Yin's diagram

The network case study model is illustrated in Figure 4.1.c alongside two illustrations reproduced from Yin (2002). Yin's original framework includes four prototypical forms: two that are single case designs and two that have multiple cases for comparison. Both single and multi-case designs can also be defined with specific units of analysis. The network model relaxes what can be seen as an implicit constraint that the sub-cases be parallel entities. It allows the sub-cases to be of different types and relate in different ways, leading to a heterogeneous case structure. Using Yin's terminology, this model variation has elements of both the single case holistic and multi-case designs. The single case involves middle school science assessment across parts of the State of

[11] In order to fit certain educational situations, I modify the original criteria that Yin used to define cases from a decision process to be a more general criterion of a process outcome. The case definition then draws on elements of Activity Theory (Engeström, 1987).

Michigan. The multiple cases involve specific teachers, schools, and a variety of organizations that support education.

Explication of the Method

Before proceeding with the description of this study's design, I include a detailed discussion of the method. This section provides additional conceptual, historical, and theoretical grounding for the methodology I am using and defines important aspects of it, including what constitutes a case.

Why a Network Model Case Study is Important

On a conceptual level, the network model incorporates a type of fidelity to underlying activities that may make it less prone to masking important details than the cellular case designs may be. Looking at the educational research process as comparable to other scientific activities (NRC, 2002), it is possible to critically look at different elements of the research design and their implications for the researcher's ability to produce certain types of warrants. In developing network aspects to the case study design, I am considering the case study design approach (for example, the embedded case structure from Young, 2006) as a type of *scientific model* that the researcher uses across the research process. Scientific models, such as the view that "matter is made up of discrete particles" or that "light is made of waves," help to explain invisible phenomena and processes in the natural sciences (Dunbar, 1993; Klahr & Dunbar, 1988). I consider them to occupy similar explanatory roles in the social sciences, although the term is not commonly used in this context.

Applying this view of the research process here, I am following Nersessian's designation of *signature objects* that help to organize the work of researchers and the initiation of new participants (Nersessian, 2007). From a representational perspective, the case structure also helps to communicate socially (Latour & Woolgar, 1979). Just as other research methods -- for example, ethnographies or statistical analyses -- prioritize some elements for analysis, so do case studies. From this perspective, case studies that use contrasts such as "privileged" and "at risk" schools (Diamond & Spillane, 2004) or “advantaged” and “challenged” districts (Falk and Drayton, 2004) may be implicitly communicating a model structure through the presentation of representative cases.

When considering a systemic study with organizational implications such as this study, the issues of case structure become increasingly important, as both the formal and informal structures of educational activities exhibit properties that are more complicated than dichotomies, continua, or even hierarchies.

Adopting a network structure in the research design represents the actual topology of educational systems more appropriately than a cellular design approach, where uniform cases are drawn from similar hierarchical levels and presented for comparison or presented as nested but essentially similar cases. The network model has another representational advantage owing to the grounding of its analytic approach in ANT. ANT prescribes a research process that does not predetermine social structure, but looks at each instantiation of social action individually for its *particular* relationships or associations. An example of this can be found in this study in the relationship between districts and schools. Typically in the educational research literature, schools are nested within districts. This study, in looking closely at the relationships around middle school science assessments, found that different districts have differing relational qualities with the schools on different topics. In some cases the district played a more significant mediating role by acting as the data processor of state test results for their schools (Firestone & Gonzales, 2007). In others, the district had a minimal role in this process, and the relationship between school and state was essentially unmediated. Additionally, from a practitioner’s perspective, some of these schools are not as nested in their districts as others are for the purposes of middle school science assessment. This study approach allows those irregularities to be more visible and allows the traditional view of school organization as a uniform hierarchy to be questioned in productive ways.[12]

Some Prior Work Supporting the Network Model Case Study

While I have not found a description of the network model case study architecture in any research literature, there are a number of previous research definitions that

[12] This study joins, then, a number of other recent works, including Spillane (2006), Burch (2006) and Coburn (2004), who focus on the non-hierarchical qualities of education in practice.

establish a solid foundation for it. Going back historically, the classic study of the governance structure of a typesetters' union during a time of technological change provides a relevant heterogeneous multi-leveled perspective using multiple embedded units of analysis (Lipset, Troow, & Coleman, 1956). More recently, Patton (2001) describes a case/sub-case structure where the sub-cases are overlapping and interrelated, but where the overall case is of a single organizational level. He separately discusses a multi-leveled case structure for a study of a state's program implementation, but the multi-leveled structure is hierarchical and nested. Significantly, Provan and Milward (1995), working with health-care systems, analyzed networks and describe the evolution of the unit of analysis in organizational studies towards more sophisticated representations:

> …most of the work in this area has focused on the determinants or predictors of inter-organizational relations (see Oliver, 1990, for a review), as an understanding of the phenomenon has grown, the unit of analysis has gradually ***shifted from the dyad to the organization set, to the network***. Especially in recent years, the study of organizational networks has proliferated. Much of this interest has been generated by an emerging recognition by academics that businesses, as well as organizations in the not-for-profit and public sectors, are increasingly turning to various forms of cooperative alliances as a way of enhancing competitiveness and effectiveness that would not be possible through the traditional governance mechanisms of market or hierarchy (Powell, 1990). (Ibid, p 1, emphasis added.)

What Provan and Brittan have done is to use the interrelated elements in an organizational network as the basis for a four-cell comparative design. This dissertation study, then, is taking a logical step in articulating explicitly the relationships between cases – as Patton (2001) did and Provan and Milward (1995) did not – into a single network design.

The network that this study evaluates -- science assessment practices across parts of the State of Michigan -- is also one that is currently undergoing significant changes, made more salient by technological innovation. This makes this study similar in some respects to the technology organizations Brown and Eisenhardt (1997) studied using case methods.

What all of these network-oriented studies, including this one, share is that a network of evidence is presented in order to reason about the larger educational enterprise. This network of evidence is an important part of the way the method supports an evidentiary process (Latour, 1987). These studies represent, I argue, an evolution of methods in response to deepening theoretical understandings about the nature of the organization as both irregular and dynamic.

The Network Model Formally Defined

In using a method that is not fully specified in the literature, it seems important to provide some definition of its key constructs and criteria. I use four essential definitions to describe the method:

1. The case is a distinct traceable phenomenon or Activity System (Engeström, 1987), rather than an abstract category.
2. A case requires multiple forms of evidence that include some historical component.
3. Cases are interrelated in the study in ways that are reflective of other interrelationships in the field being studied.
4. The empirical components are arranged in an *evidentiary web* that is an object of validity inquiry.

The first characteristic is that a case is a distinct social process that has left a trace. A trace is usually a text of some kind that indicates past action and/or future activity. A general category, such as "teachers" or "districts" could not then be cases, but professional community organizations that represent teachers and districts and include them in various participation structures could be considered cases. These professional groups are not proxies, however, and using them entails accounting for their representational limitations. While Yin (2002) uses the decision as a key criterion of a case, I relax this constraint to adopt the outcome aspect of activity systems (Engeström, 1987) instead. This change is helpful in cases where a process is not always clearly oriented towards a decision, such as professional associations that may not be specifically focused on decisions, but are organized also for community purposes.

The second characteristic is that the case will have multiple forms of evidence. A teacher's classroom assessment practice is a type of case for this study, but not all teachers who described their practice were used to build cases. Only those with multiple data sources (and for these teachers there were also several data collection events) were used to develop a case with all of the recordkeeping and validation processes that are associated with individual cases. This methodological constraint also entails the grouping of evidence elements that are not full cases into cases, as discussed shortly.

The third characteristic is that cases are interrelated in a way that is reflective of what naturally occurs in practice. Rather than unique or coincidental relationships, the cases are related in specific ways to other cases in the study, and these relationships are hypothesized to represent other relationships that exist in the field being studied. There are many ways that cases can be interrelated.

The fourth characteristic is that empirical components create an *evidentiary web*. An evidentiary web refers to the collection of cases as a sample that is representative of the larger network of entities that exist in reality. Much in the same way that a sample in quantitative research has representative capabilities and can be evaluated in terms of various biases and opportunities for random error, so is the evidentiary web a key component in the research process and the object of validity investigations (Messick, 1989).[13]. For example, if the cases in the evidentiary web represent certain types of entities and not others, or certain types of relationships and not others, then its overall evidentiary ability will be limited by what it includes. The combination of entities and relationships in the evidentiary web also may provide opportunities for member checking, pattern matching, and searching for confirming and disconfirming evidence.

Granularity and Grouping: Two Critical Issues

In relaxing the implicit constraint on uniformity of cases in multi-case designs, two issues now emerge as important. These issues of granularity and grouping were salient in more traditional case methods, but might have been referred to generally as "unit of analysis issues." In this study, the result of this heterogeneous design is a greater number of cases than might be present in a more traditional approach. And it becomes

[13] I am using Messick's reference here to focus on the investigative part of his approach to validity. He did develop formulaic descriptions of validity, but framed them within a process of scientific inquiry.

clear upon some reflection on the design that some of weaknesses of the cellular design approach actually have been transferred to a smaller grain size rather than excised from the design. The decision to create many cases for a school or a district rather than one still involves design choices with trade-offs. The finer resolution does not eliminate the lines that researchers draw around data, and the case is ultimately an abstract grouping according to design principles (Simon, 1996).

Further, the constraints I have imposed on cases in terms of the data they should have (multiple sources of evidence over time) means that there is important information that does not rise to the level of a case, but that should be used in the study, and it is shown within case reports. To address this issue I will join pieces of data, including interviews, from non-case participants into another case so that the information can still contribute to the study. As a researcher, I also have the flexibility to take some of this subsumed data and develop another case with additional data collection, so that the case structure is malleable: with additional research, new cases based upon information grouped into an existing case are possible.

Context: the Overlapping Formal Structure in Michigan

Michigan is a state that challenges the characterization of a hierarchical educational system. In looking across the state, there are more than 800 districts of varying sizes as shown in Figure 4.2[14]. On the small end of the spectrum are those with several hundred students. At the other end is a single urban district with over 130,000 students and the type of hierarchical structure more commonly discussed in studies of school organization.

Some of the variation across the state where the population is very dense in a few urban centers, and largely rural across most of the state where the districts tend to be small as illustrated in Figure 4.3[15]. Operating between the district and the state is a layer

[14] These data extracted from the National Center for Educational Statistics, dataset on schools and districts report number NCES-425072459 for the 2004-2005 school year. The general shape of this distribution for this one state is very similar to the shape of the distribution of national data using the same intervals although the national data contains no empty intervals and has districts of much greater size.

[15] Image produces from combination of U.S. census map showing population density, Michigan Association of Intermediate School District Administrators (MAISA) website image and Math/Science Support Center Network website image.

of 57 intermediate school districts[16] (ISDs) can be seen in Figure 4.3.a. These ISDs are also diverse. They are mostly governed by elected bodies and sensitive to the economic and demographic characteristics of the schools they serve. The ISD organization (Figure 4.3.c) is similar to the county structure, although the relationship is not parallel. Figure 4.3.b shows a typical example of school districts within an ISD along with the boundary of the county. While the county follows a fairly regular rectangular boundary, the school districts are a patchwork and as seem to cross county boundaries as often as they remain within them. The service relationship between districts and ISDs varies as well. In some cases, including within ISDs relevant to school in this study, several intermediate districts cooperate on programs so that students requiring special services may or may not have received those services from their own ISD.

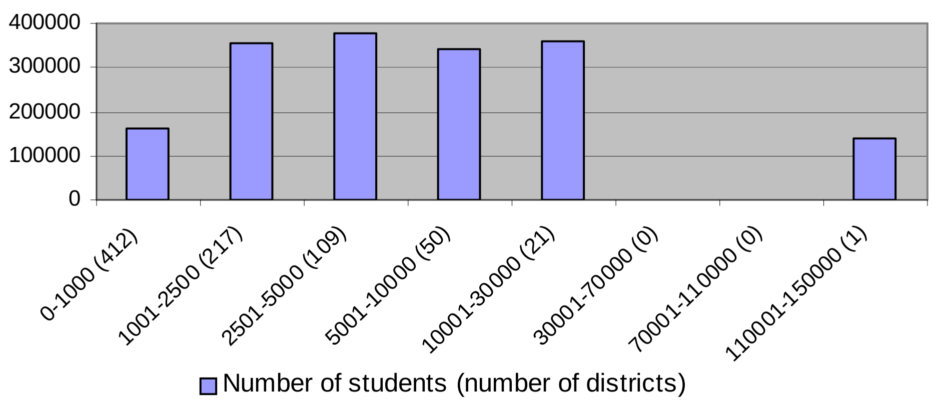


Figure 4.2 - Distribution of students into districts of different size in Michigan

Another semi-hierarchical layer in Michigan is the network of 33 Math/Science centers (Figure 4.3.d) that provide special services including workshops and in some cases science materials to schools. In some cases these centers are located within ISDs. In others, a center is shared by more than one ISD. A few are housed within specific districts and three centers are part of university programs. Some centers provide a few services and are little more than a single office and staff person while one center, the Battle Creek Math Science Center, occupies a former high school and provides science material kits to more than 25% of the students in the state (Connie Duncan, Personal Communication, February 12, 2007).

[16] These aspects of Michigan have received some attention from Spillane (1998) and Cohen (1995).

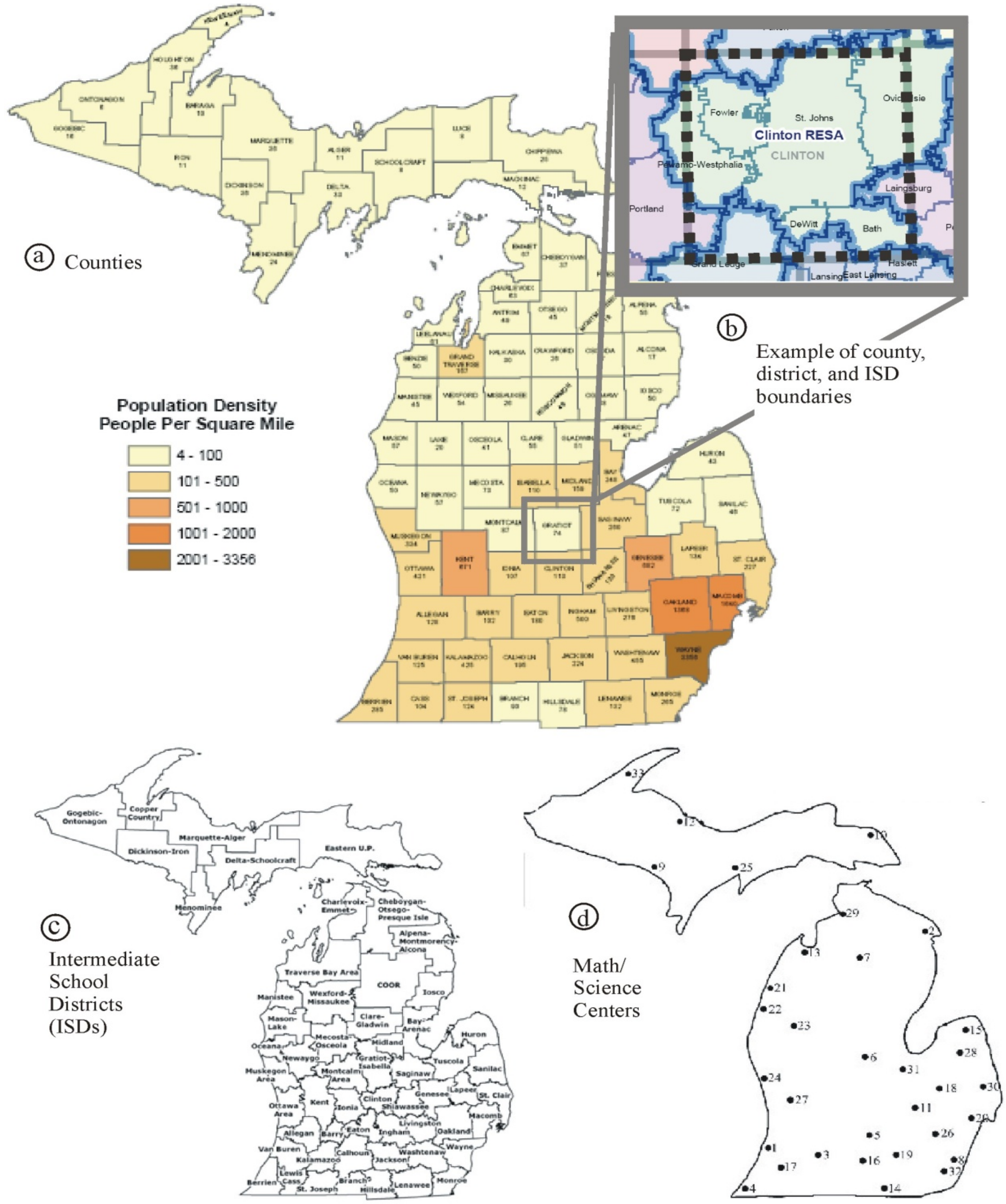


Figure 4.3 - Characteristics and organizational structures in Michigan

At the time this study was conducted, Michigan was in a period of general economic recession with many industries, including automobile manufacturing and its network of sub-industries moving jobs overseas. The state budget and the investment in education had both been significantly impacted by these conditions with lower available funding for education resulting. The state also supported some measure of school competitiveness so that parents were able to select from different public school districts they are close to. As a result, many districts were facing pressures to compete for students and funding as they had not in the past. Further, there were efforts to consolidate various districts making the level above schools not only irregular but also dynamic and changing. One area that was under intense pressure was the network of Math/Science Centers. They were established in the early part of the 21st century with federal and state grants but had faced budget cuts such that some were supporting themselves through fees or as is the case of Battle Creek's center by acting similar to a commercial publisher. The state government was also competing for the scarce education funds in the state by selling various services including a school improvement plan that bears some similarity to the Data Wise (Boudett, City, and Murnane, 2006) approach reviewed in Chapter 2.

The system of testing for NCLB in Michigan includes annual assessments of math and literacy but science was only assessed in 5^{th,} 8th, and 11th grades. The elementary and middle grade tests were called the Michigan Educational Assessment Program (MEAP) and the high school tests fell under the Michigan Merit Exam (MME). These are similar programs that were until the year of my study both called the MEAP. The MEAP was developed primarily by the state with contractor support while the MME was built using the non-profit organization ACT's framework and with Michigan government and practitioner support. A perceived lack of coherence between the MEAP and the MME had been cited by a number of this study's participants. Also, since this study looks at the 8th grade test which assesses what students should have learned in 5th through 7th grades it showed the last opportunity for schools to use this type to assessment information before the students entered high school.

In the area of science standards and testing, the state was undergoing important changes as well. The standards framework in use in 2007 was from 2000 and a new set of science benchmarks called grade level content expectations had recently been developed by the state with input from Michigan science education experts. These new standards were scheduled for public release in 2008 and should then have been available for use in student testing in 2009. However, given that the test development process operated on a two-year timescale, the MEAP development process was not aligned with the new frameworks because those new frameworks had not yet been officially approved and could have change in a review process. This lack of coordination between standards and test development added to general uncertainty about science education in the state during this period of this study. Some of those who worked with the state to develop GLCEs and MEAP/MME test components participated in this study.

A Brief History of the Study

To introduce the sample, the evidentiary web for this study, I will first present a brief history of the study itself to explicate some of the sampling choices I made. Many of the characteristics of the sample used in this dissertation are artifacts of the study process itself. Others reflect participation choices by different districts and sites where access was requested.

The study did not begin as a network case study that was exploring different organizational levels in the State of Michigan and attempting to make inferences about various systems working in the state. Rather, it was initially conceived of as a *school-based* study that would compare several schools in different types of districts, in different states, to develop a common model of the perspectives of teachers, principals, and curriculum specialists, what they would have in common and how they would differ in their approach to assessments.   It was initially a study seeking to develop a *general model* of educational organizations related to science education in middle schools.

The methodology originally proposed was a variation of ethnography called *textography* (Swales, 1998) that is designed to investigate the roles that texts play in local discourse communities. The dissertation proposal was theoretically grounded in a combination of Weick's model of educational organizations as loosely coupled systems (Orton & Weick, 1990; Weick, 1976) and Lemke's ecosocial timescales model (Lemke & Sabelli, 2006; Lemke, 2000).

The study began in the fall of 2006 with active school recruitment both inside and outside of Michigan. The first schools to agree to participate were Michigan schools, and the initial findings from these schools indicated a general paucity of science assessment resources and scarce organizational attention to middle school science in comparison to other subjects in other grades. From an early point in discussions with Michigan schools, it appeared that classroom assessments were mostly localized to an individual teacher's practice and unlikely to cross organizational boundaries, unlike what the literature reviewed in Chapter 2 indicates is possible in other areas, such as literacy in primary grades. The annual state assessment results constituted the principal alternative to these individual classroom assessment systems. As a result, I considered extending the study to include a focus on the state testing operation, the Michigan Department of Education's Office of Educational Assessment and Accountability (OEAA). By early 2007, the study had changed from a comparison of schools across several states to a study that focused on the State of Michigan at both the school and state level.

This realignment in research focus was further supported by the timing of the release of the state's test results to the schools in February of 2007, which created some opportunity to study the brief process by which some schools make sense of the state results. In addition to adding new types of data and research sites to the study, the inclusion of the state testing office meant that the State of Michigan became a primary unit of analysis, and making inferences about what was occurring more broadly across this particular state became important as well.

The combination of these developments then changed the nature of the study from one that was intending to develop a general organizational model of how assessment might operate in middle schools to one that is particular in a number of important dimensions. It is particular to the State of Michigan and its decentralized governance structure. It is also particular in some ways to those organizations that provided access during the period of data collection. And as data collection proceeded, it revealed itself to be historically situated in important ways. As the study progressed, this historical particularity became a major theme undergirding the entire project.

The restructuring of the study from its original investigation of schools to a study of a state also brought in issues of districts and support centers that operate between the schools and the state. However, access to these intermediate structures was difficult in much the same way as access to some schools was difficult. Anecdotal evidence at the time indicated that middle school science assessment was an area that these intermediate support structures had also barely begun to address, and the public exposure revealing these educators' low preparedness might lead them to avoid participation. Still, these intermediate structures are important parts of the state system and some evidence related to them seemed important for the study to be complete. It was the difficulty in studying these intermediate structures that led me to decide to study *professional associations,* rather than study them through research sites in the way the schools and OEAA were studied. Because these professional associations produce public texts, it became possible to create cases for them.

As a result of this particularization and the new scope of the study to include locations other than schools, the basic methodology was shifted, as well, to the network case-study method. This new method introduced both new theoretical elements as well as new requirements for the data. As I made decisions about what to consider as distinct cases, I was then pushed to gather new evidence to strengthen those cases and to look closely at case boundaries and interrelationships. One important additional source of data this shift led to was a survey of the membership in the Michigan Science Teachers Association that acts as an important resource for comparison with what was found in school cases.

The Meta-Model and Overview of the Sample in the State of Michigan

Since the network model introduces heterogeneous case structures, there are some important considerations and options for presenting the case interrelationships that may not be as significant in cellular case-study designs. Drawing off of information engineering, I will present the inter-case structure in two parts: meta-model, which describes types of cases and inter-case relationships; and a description of the sample in this dissertation -- the dissertation's evidentiary web.

The Case Meta-Model

The case meta-model defines the rules (instantiation possibilities) that may exist in the actual data. In this study, it broadly describes three classes of organizational cases, along with an individual case type that may participate in more than one organizational case, as illustrated in Figure 4.4. The lines on this meta-model indicate potential inter-case relationships. Typically, individuals are grouped within an organization. But in the domain this study explored, some individuals play significant boundary roles. This qualitative model is sensitive the boundary activities that they engage in relative to assessment, and the meta-model supports this by separating the individuals from organizations and then relating the individuals to those specific organizations in the study of which they are a part.

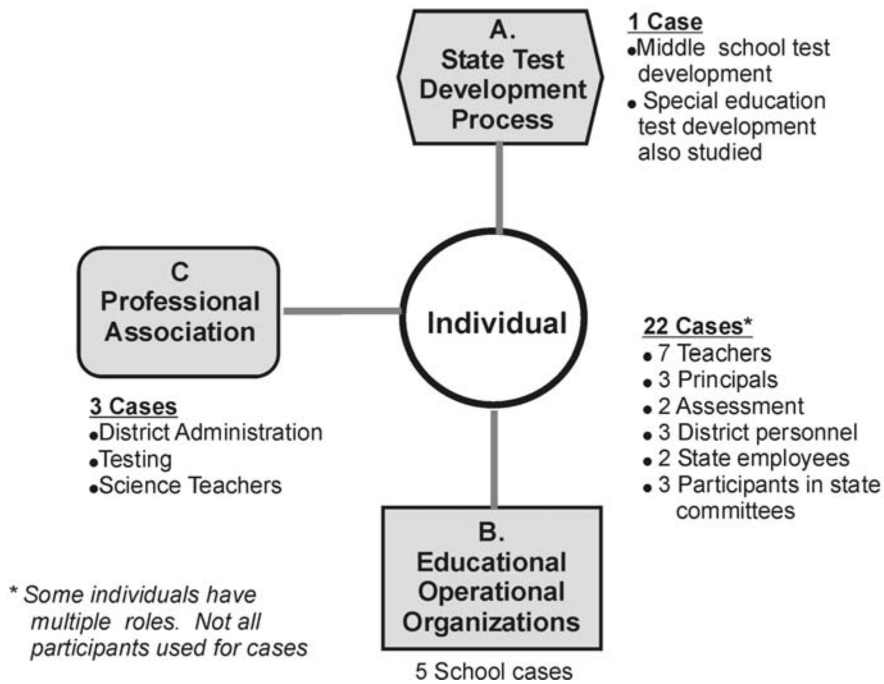


Figure 4.4 - Study data meta-model schematic

The three major categories of organizational cases serve different functions in the study. The state testing office (Group A) represents the highest level of analysis. The study's foundation is the set of cases called Educational Operational cases (Group B). These are organizations with day-to-day responsibility for students and instruction. The core of this group is made up of schools and teachers, with some information from local and intermediate school districts (ISDs). The inclusion of districts in the operational group, rather than as a separate administrative group, as they do have a governance function that is distinct from school operation and the state, was made because that is an area where the study's data was often insufficient to develop full cases. And, in the area of middle-school science, the district relationship to practice was often weak. Generally, the specific district and ISD information that was collected in this study was grouped into the school cases.

The professional association cases (Group C) serve three principal roles. They provide a framework for interpreting what is found in the operational cases. The professional associations provide first evidence over longer timescales (Lemke, 2000) than the evidence collected in the operational contexts, so it is possible to relate local evidence to broader community patterns. In this way, it is possible to evaluate what was observed in schools as either typical or extreme. Second, they provide a way to extend inferences to larger groups of schools across the state. In a sense, they serve as a cross-case triangulation source to amplify, where possible, what the data in the foundational operational cases show. And third, they provide evidence about areas of the system where there is little observational data collected.

Overview of the Cases; the Evidentiary Web

The different cases that form the evidentiary web for this work will be discussed in more detail in the analysis chapters that follow this one. The five school cases were drawn from five districts of different sizes from each of the first five size categories shown earlier. At the state level, there is one case for the state's OEAA unit responsible for middle school science for students without severe disabilities. This case is central to the research goals of this dissertation, and its documents, observations of meetings, and interviews with its key personnel are used in cross-case analysis for several parts of the study.

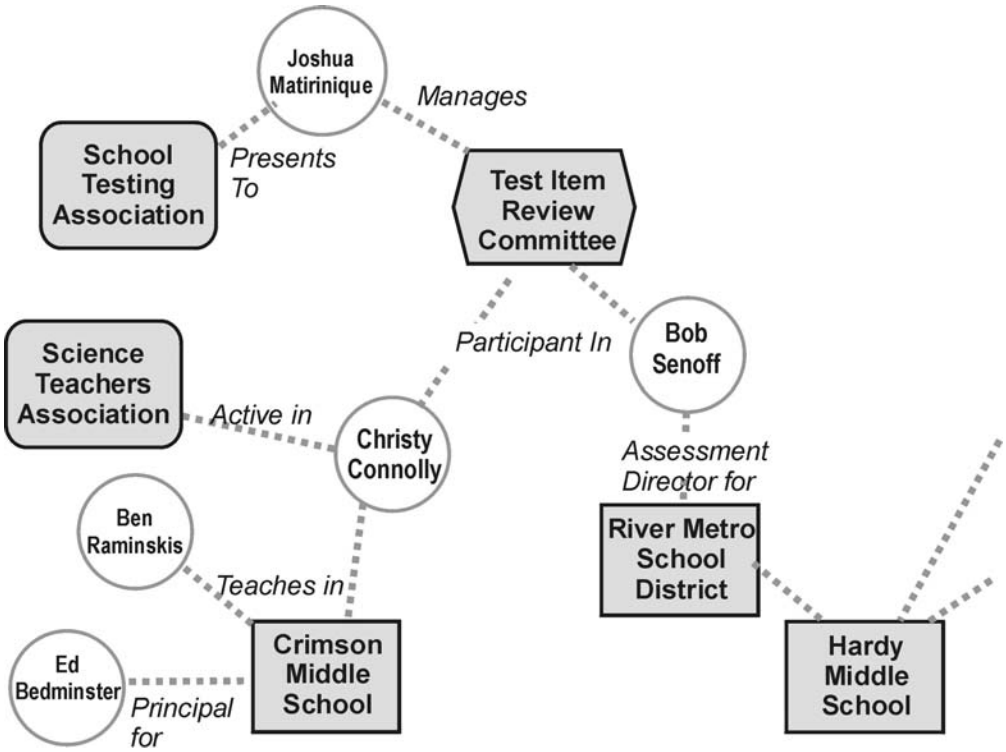


Figure 4.5 - Network model case structures

One of the key boundary spanners in this study is teacher Christy Connolly.[17] She is a classroom teacher who is active in the MSTA and also participates in the state test development operation. She is also a user of a classroom technology called Moodle™ that is important for her assessment practice. In order to understand her practice, I contacted the local district's technology person responsible for Moodle, Annie VanDusen, to understand the nature of both the initiative at the district and Christy's work. If enough data had been collected for the district, it could have been considered a case. Since the district is not recorded as a case, then this extra bit of data that helps me to understand this teacher's work needed to be categorized in some way. In some cases, as with this interview with Annie VanDusen, the data is cataloged with the case records of the teacher Christy. As it turns out, while I do have district data in a number of cases, I have chosen not to have any district cases and instead collapse that data generally into school cases.

[17] All individual names are pseudonyms

Table 4.1 - Overview of cases

Category	Case Name	Description	Number of participants
A. State	MEAP Development	State test development process	9
C. Professional Association	Assessment Community	Mich. Student Testing Conference/ Mich. Educational Research Association	4
	School Admin. Community	Association of School Administrators	5
	Science Teachers Community	Michigan Science Teachers Association	5
B. Educational Operational	Avon Falls	Middle School	6
	Crimson	Suburban Middle School	5
	Dorchester	Affluent Rural School	2
	Hardy	Large Urban Middle School	7
	Swallow	Small Middle School	5
Individuals	Ben Raminskis	Classroom teacher	1
	Bernie Lauer	Assessment specialist	3
	Betsy Dearing	Classroom teacher	1
	Bob EnSpania	District Superintendent	1
	Bob Senoff	Head of assessment and evaluation	1
	Burt Wainwright	Principal and curriculum coordinator	2
	Christy Connolly	Classroom teacher	4
	Corrine Eaton	Science Education Leader	2
	Dianne VanderMiller	Coach and assistant principal	3
	Don Pulte	Assistant Superintendent	2
	Faith Churchill	Classroom teacher	2
	Hardy Kit Teacher	Kit teacher at Hardy MS	5
	Janey Fess	Curriculum Director	1
	Jim Heinrich	Classroom teacher	1
	Joshua Martinique	Manager of state assessment	1
	Karen Minor	Curriculum Director	2
	Nancy Newman	Curriculum Director	1
	Paul Bond	Classroom teacher	3
	Pete Darmond	Science Education Leader	1
	Roscoe Ellis	MEAP Science Consultant	4
	Stan Dubovski	Principal at Swallow Middle School	1
Other	Successline	Software Vendor	3

Data Categories

This study has five different categories of evidence: interviews/personal reports; fieldnotes; observations; documents (texts); and surveys/anonymous structured instruments, as shown in Table 4.2 below. Within some of these categories are sub-categories that are associated with some of the individual cases. As the cases are presented in the analysis chapters, the evidence available for each will be discussed. as well as my observations about the nature of those data related to the case.

Table 4.2 - Data types by case category

Evidence Base	**Cases Within Category**			**Types of Evidence**				
Case Categories	**Organizational**	**Individual**	**Non-case informants**	**Public Documents**	**Internal Documents**	**Surveys**	**Observations***	**Interviews***
A. State test development process	1	2	26	X	X		X	X
B. Educational operational (schools/districts)	5	15	14		X		X	X
C_1. Science teacher professional association	1	2	4	X		X		X
C_2. School testing/evaluation professional association	1	1	5	X			X	X
C_3. School district professional association	1	2	5	X		X		X

* - Recorded with audio, video, and in some cases fieldnotes

Interviews/Personal Reports: Individualized and Common

Personal reports in this research consist of both individualized and common elements. Both were collected in interviews and in almost all cases, recorded with either audio or video. For the individualized interviews, the participant was queried about a topic from either a developed interview guide or as a conversational question in the interview. The common elements consist of a series of questions that were asked similarly of most of the participants. Appendix B contains a complete list of questions asked in this study.

Field Notes and Meeting Notes

Fieldnotes are artifacts that I made either within a research event or, in some cases directly thereafter to record my observations that would not have been recorded otherwise. When I was able to use audio/video recording, I generally did not also take fieldnotes. In some cases, consistent with ethnographic techniques (Emerson, Fretz, & Shaw, 1995), the fieldnotes were re-written into typed form, and in some cases they were captured as an image from my note pad. There was one participant, the teacher John Heinrich, who did not want to be recorded, and so all interviews with him were recorded in notes and notes were used to record a department meeting at which he participated. Fieldnotes were also used to make records of some telephone conversations.

Meeting Observations

Several meetings at the state and school level were directly observed and recorded. These meetings, three at the state and three in schools, were focused around the MEAP assessment items. The meetings at the state involved the review of items developed by teachers within the state-defined assessment item development process. The meetings in the schools involved the review of that school's or district's results. While the review process in the three schools studied was not as formal as the state item development process, it was possible with each of these local events to develop a sketch of the larger organizational processes within which the meeting was situated.

7th Grade Science Quiz #2 Name ___________
Matching; Hr_____

_____ 1. nucleus	a. a tool for looking at cells
_____ 2. cell membrane	b. a storage room in a cell
_____ 3. cell wall	c. the cafeteria or food maker in a plant cell
_____ 4. cytoplasm	d. the boss or master controller of the cell
_____ 5. mitochondria	e. a glass plate for putting objects on to be looked at under a microscope
_____ 6. vacuoles	f. the rigid structure on the outside of a plant cell or a building
_____ 7. chloroplasts	g. the security guards of a cell that let certain things in and out
_____ 8. microscope	h. the thin plastic cover you place over an object on a slide
_____ 9. slide	i. the power plant of a cell where energy is produced
_____ 10. coverslip	j. the liquid like part that delivers nutrients to all parts of the cell

Extra Credit; Who discovered cells in 1665? ______________________________

Document 4.1 - A quiz as an example of a private document

Documents

Documents form one of the most common types of data in this study, since they are part of every case. Quite often, each case type entails specific document types. For example, teacher cases have examples of classroom instruments, such as quizzes and journals, and professional associations have newsletters and conference proceedings. In some cases, the documents collected would be unlikely to receive much notice outside of an activity system. For example, an individual teacher's test forms are usually limited to the classroom and maybe students' homes. More public documents are often available on websites and are possibly archived documents that could be available for some time after this research was completed. Documents 4.1 and 4.2 show some examples of documents. The analysis chapters will provide a description and images of documents in closer proximity to their use as evidence.

Michigan Curriculum Framework Science Content Benchmarks Summer

All students will explain what the world around us is made of:

(No elementary benchmark about molecules or atoms)

3. Classify substances as elements, compounds, or mixtures and justify classifications in terms of atoms and molecules.

Key concepts: Element, compound, mixture, molecule, atom. See PME-IV.1 m.4 (molecular structure of solids, liquids and gases).

Real-world contexts: Common substances such as those listed above, including—elements, such as copper, aluminum, sulfur, helium, iron; compounds, such as water, salt, sugar, carbon dioxide; mixtures, such as soil, salt and pepper, salt water, air.

4. Describe the arrangement and motion of molecules in solids, liquids, and gases.

Key concepts: Arrangement—regular pattern, random. Distance between molecules—closely packed, separated. Molecular motion—vibrating, bumping together, moving freely. (PCM-IV.2 m.4 addresses the molecular explanations of changes of state.)

Real-world contexts: Common solids, liquids, and gases, such as those listed above.

3. Explain how elements differ, in terms of the structural parts and electrical charges of atoms.

Key concepts: Parts of atoms—nucleus, electron cloud. Subatomic particles—proton, neutron, electron. Electrical charges—positive, negative, neutral. Each element has a unique number of protons. See PMO-IV.3 m.3 (electric force).

Real-world contexts: All elements.

Document 4.2 - Selection from the 2000 Michigan Science Standards as an example of a public document

Surveys and Anonymous Structured Response Data Sources

There are several data sources where individual responses are structured and are used to provide important information localized to cases. Some of these instruments have been called surveys, but may not meet many expectations of surveys. Each of these data sources will be used within specific analyses and generally as supporting or triangulating evidence, rather than in a primary role. Within each analysis chapter, the specific ways that each data source is used will be described.

Table 4.3 - Surveys and structured anonymous data sources

Instrument	Created By	Timeframe	Respondents
3 ISD member surveys of data warehousing and common calendars	School administrator associations	DW: 2005/2007 Common calendar: 2006	Majority of ISDs
1 Survey of science professionals' assessment use	By me for the dissertation	August-October 2007	198 members of Michigan Science Teachers Assoc.
1 Focus group questionnaire	By me for the dissertation	April, 2007	6 committee members
1 Public review of MEAP items	OEAA	March, 2007	Educators at public review meetings

ISD Data Warehousing and Common Calendar Surveys

The Michigan Association of Intermediate School Administrators, the ISD Association, conducted two surveys of their membership's use of data warehouse technology. One was conducted in 2005 and the other in 2007. Thus far, I have only been able to able to gain access to the executive summaries of these documents and am working to get the raw data. They portray the contemporary emergence of the types of infrastructures discussed in many of those sources reviewed in Chapter 2. There is also a survey of ISD efforts to develop common calendars, conducted in 2006, that provides further evidence for the period of standardization and structuring that is now being attempted across this state. Most ISDs (> 40) participated in all of these studies.

Science Professionals' Experience and Beliefs about Assessment

In order to broaden the types of claims that this study can make from the five schools and seven teachers that were interviewed, I developed a survey that was sponsored by the Michigan Science Teachers Association. This survey is discussed in Chapter 6.

Questions were asked about classroom assessment practices and the Boundary Practices (Wenger, 2000) of reviewing MEAP information in schools and participating in the MEAP development activities. The respondents to this survey were mostly teachers and those who work in schools from districts of many different sizes, so they represent an opportunity to look for confirming/disconfirming evidence for patterns and trends suggested in other parts of the study.

Focus Group Questionnaire

The focus group consisted of members of a MEAP review committee who were asked some of the same questions asked of the case study participants. This questionnaire included a written portion as well as a recorded set of questions. The idea of using a review committee as a focus group came from the Michigan Department of Education, which made time in their meeting for my data collection.

Public Reviews of MEAP Items

As part of the development process for MEAP items, there are general review sessions where educators can look at questions under development and make comments that are then entered into a database maintained by the Michigan Department of Education (MDE). MDE gave me a copy of these review files, and I used them in a small way for pattern checking and further confirmation of the item's role as a quasi-boundary object.

Analysis Process

The analysis process in this study generally follows a case-based pattern of building an evidence base in stages with attention to validity. The process began with two concurrent activities of indexing the data: thematic coding and developing case report documents, followed by member checking and cross-case synthetic analyses, to be followed by three semi-independent analyses. Through this sequence of analytic steps, the chain of evidence was constructed, beginning with raw data that was coded and interpreted by me and then validated as case reports with the participants. Each of these activities is discussed in more detail below.

Data Indexing and Thematic Coding

Data indexing and thematic coding involves reviewing the recorded media and fieldnotes and assigning qualitative codes to sections of the media that relate to different theoretical aspects of the dissertation. The purpose of the indexing and coding was to complete the data inventory and allow me to locate any given response by a participant that related to an analytic theme.   This coding scheme (discussed below) is hierarchical and emergent. It is hierarchical in that it involves codes and sub codes. However, any given data element, for example a participant response, could be multiply coded if it were addressing more than one topical aspect.

Development of Case Reports

The case reports are short narratives that describe the nature of what this study observed. The case reports were written in a conversational style and then shared with the participants who helped inform the cases. The case report writing will structure the data coding process, as I will proceed in a case-by-case basis to review the data I have and catalog all documents and data collection events concurrently with the work of the case reports. This process will be aided by my analysis database, which has records of each meeting and event and indexes the artifacts that come from each. Each case report will include, at a minimum, three main sections: a short overview, case details, and a summary with questions that remain for me. This summary will be my way of communicating an overall impression of the case. Each case report will include some common elements, including:

- A narrative that discusses what the case entails, for example, detailing the assessment practices of a given teacher.
- A short description of each data collection event.
- A short summary of all data used in the case.
- A list of related cases.
- A list of all individuals related to case.

Qualitative Coding Structure

Part of my analysis uses a set of qualitative codes to index and categorize the various data components (ex: segments of talk). Some of these codes identify broad topical areas, such as when an individual is describing her background. Others are specific to an assertion, for example, that accountability testing is a positive or a negative. I have structured these codes into a two-level group/category structure. Within a given group, identified by a two-character code, there may be multiple individual category codes. Table 4.4 shows the high-level group codes divided into three sets. One set relates to markers used to index study components but not used in analysis. The second set involves single level code groups, where the group does not differentiate between different ecosocial levels. The third group, called system topic codes, do span ecosocial levels and can be further qualified by a one-character ecosocial level code.

This is an emergent coding scheme that changed and grew through the analysis processes. In the coding system used in this dissertation (structured by the analysis software) the sub-code follows the topical code and a dash. For example, epistemology includes sub-categories that "not all children can succeed" (EP-NK) and that "all children must succeed" (EP-AK). This coding structure allows both a fine level classification of evidence and the ability to make higher level comparisons.

Table 4.4 - High level qualitative codes

Code	Title	Description
Codes not used in content analysis		
PI	Personal Information	Information about a participant. Includes demographics.
EV	Evidence Base	These are notes related to the evidence base of the study. Used to document cross-case linkages.
Single level codes		
PO	Policy	Policy references. Could conceptually be considered activity system reference at a super-ordinate level.
ST	Students	Reference to students either specifically or as types of actors. These are always classroom/local references
EP	Epistemology	An individual's opinion: that testing is contrary to good instruction, for example.
Systemic topic codes		
AA	Assessment Artifacts	References to specific characteristics of assessment artifacts.
AI	Activity Interactions	References to interactions between what can be considered activity systems (ex: school-home or school-district).
AS	Activity System Descriptions	Description of aspects of an activity system in the study, for example a teacher's classroom practices or the state standards.
HC	Historical Change	Reference to how things have changed from the perspective of a participant.

For the systemic topic codes, the optional ecosocial qualifier differentiates the level that is being referenced in the specific element of evidence. For example, within the systemic topic of Activity System Descriptions, I have further subdivided major categories of state, local and classroom. The codes then make it possible to differentiate between evidence that provides a description of the state assessment practice (ASS) and a classroom assessment practice (ASC). Table 4.5 shows the multi-level topic codes.

Table 4.5 - Multileveled coding structure

----- Systemic Topics -----				
Historical change (timeframe)	**Activity structure & Practice**	**Activity Interactions (coupling)**	**Assessment artifacts**	**Systemic aperture/ Ecosocial level**
HCS	ASS*	IAS	AAS+	**State (MEAP)**
HCL	ASL	IAL*	-	**Local (School/District)**
HCC	ASC*	IAC	AAC+	**Classroom**
-		IAX	-	**External to schools**

* Timescale component,
\+ Representational characteristics

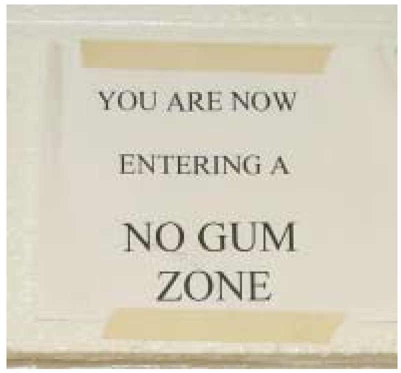
Research Support Technology

The study uses some central definitions that are embedded within the technology, consisting of a case records database and a transcription system. The case records database was developed by me to maintain the records of this research and is implemented in Microsoft Access operating on a Windows XP operating system. The transcription system is the Transana product developed by the University of Wisconsin (http://www.transana.org/). The Transana system has a robust definition of transcription and can be used as an indexing system where locations in a media file (audio or video) are saved in a file that be used to both assign qualitative codes and to navigate to the original locations for further analysis in addition to the traditional written record of speech and gesture (Ochs, 1979).

Transcription Conventions

The transcripts presented in the analysis of this dissertation employ a basic convention where each conversational turn is assigned a number. Latching and overlapping speech is not shown, nor is prosody indicated. I use two types of special indications for emphasis. Capitalized text indicates emphasis by the speaker, as indicated by an increase in volume. Transcript 5.1 shows an example of this kind of speaker emphasis. Bold was used to indicate emphasis by a change in pitch, but not an increase in volume. Transcript 5.3 shows an example of this emphasis. Italics are used for re-voicing, where the change in speaker intonation indicated he was relaying what someone else or he had spoken of on a different occasion. Transcript 5.4 shows an example of re-voicing.

Chapter 5 Classroom Assessment Activity



A No Gum Zone

Hanging over the entrance to Faith Churchill's science classroom in Dorchester Middle School is a sign stating "YOU ARE NOW ENTERING A NO GUM ZONE." Students would find it hard to miss this sign as they entered her classroom. Faith is one of seven seventh-grade teachers who participated in this study -- a study that begins by describing how these teachers approach assessing what their students have learned and what they do with the information from those assessments.

This dissertation looks at a single topic – seventh grade science assessment – across a number of organizational settings in a single state. It is both a science education study and a study concerned with issues of educational organizations and how they interrelate. From the comfortable distance that is usually afforded organizational studies, grades and subjects can appear with a grid-like simplicity. From that view, the seventh grade is one year later than the sixth and the year before the eighth. The teachers in this study, however, provide another picture. They all told me that up close, seventh graders look and act different from sixth and eighth graders. This grade that sits in the middle of what is most often called middle school (formerly junior high) is a time of important growth and transition for emerging adults. As the case studies that follow show, teaching and testing these students often involves a mix of viewpoints and techniques.

According to Jim Heinrich, another study teacher, seventh graders still retain some elements of little children. Betsy Dearing, who teaches in the room next to Jim, described these years as the kids doing “that middle school thing.” One day early in my study, Betsy and I were on our way to discuss her approach to classroom assessment. Passing through the office, we were delayed while she stopped to administer first aid to a student who limped in holding his foot, grimacing. Later, when I asked her what she meant by “the middle school thing,” she replied:

> Well part of being a teacher is – like you witnessed – that kid has a cut on his ankle in the office and why did that occur? And more the social aspects ... those things don't get covered on tests... y’know...the best friends are you know *enemies* today and why that's going on um how those kids are doing ...especially in the middle school, you have a lot of issues taking place with boyfriends, girlfriends, he broke up with me, I'm jealous of you. Friends.
>
> And you know the disorganization is a big one in the middle school. Things that have worked for years, you hit sixth, seventh, eighth grade and the kids just can't keep track of anything anymore. They are so worried about their self image and image to their peers, they just lose it all. (Dearing, personal communication, February 23, 2007).

Jim said that part of the challenge for seventh graders is that they are beginning to learn new types of concepts. Faith Churchill also articulated that view, saying that until seventh grade much of the learning involved memorization, whereas once they get into her class, students begin dealing with more integrated sets of ideas. How much of this transition is a matter of student cognition and how much of the explanation lies elsewhere is beyond the scope of this study. Christy Connolly, however, divides her classroom, with boys on one side and girls on the other, to create an environment where they can feel comfortable exploring the natural world without the peer and social pressures that might otherwise get in the way of their enjoying science. As Valerie Jones, who teaches in one of the state’s largest middle schools, said of her students, “They are very social, and science is not a top priority unless it is about them.”

As the year progresses, the students change. Valerie Jones said that it is not uncommon for some kids to have crying episodes at the beginning of seventh grade, but that by Christmas those episodes are finished. Ben Raminskis, who teaches next door to Christy, finds them "pretty eager to learn ... just more excited about science than the older kids." He is one of two teachers in this study who teaches both seventh and eighth grade, and he said that "as they get down to eighth grade they tend to get a little more jaded ... a little harder to keep in focus. Eighth graders also have less of the 'he took my pencil' foolishness than with the younger kids." Paul Bond, who teaches in this study's smallest school, said that the kids come from a highly structured environment in sixth grade, and he spends the first half of the year giving the kids a lot of support and guidance until they are more comfortable with all of the demands of changing classes and having greater responsibility than they have had before.

Whatever that *middle school thing* is, the seventh grade can be its exposition. While the seven experienced teachers in this study focused on different aspects of the age, they generally agreed that it was an important transitional period and that seventh grade was consequential for their students' futures. All of the teachers stated a belief that, for at least some of their students, seventh grade science could mean something for them later on. I begin this chapter with glimpses into what these teachers experience daily to reinforce the idea that, while the analytic frame in this research is systemic, at the classroom level the system is largely concerned with what Cohen calls the business of human improvement (Cohen, 1995). In other words, if the work these seven teachers do can be considered *a system*, it is one that involves fundamentally human issues in which child and young adult can inhabit the same body.

Introduction

The goal of this chapter is to describe the classroom assessment practices of seven teachers from five schools from a systemic perspective. This chapter begins with case studies for these seven teachers, followed by a series of short cross-case comparison analyses. These analyses treat issues of the systemic composition of these different teachers' praxis, including the different types of instruments they use, the types of semiotic tasks students are given, the temporal structuring of their assessment approaches, and some of the perceptions and dispositions of these teachers regarding systemic changes and accountability. Each of these analyses presents a facet of the systemic nature of the assessment practices of the seven particular teachers and helps to develop a depiction of some of the variation in teacher practice. I will use this depiction, as well as refer to many of these teachers, in later chapters as the perspective of this dissertation broadens.

Seven Teacher Cases

The seven cases that follow each have a similar outline derived from the full case reports produced by the study. While most of the case report information is presented in the following seven sections, some parts appear in the analysis portion of this chapter, and others are only summarized.

The Case of Paul Bond in Swallow Middle School

Paul Bond (PB) teaches both seventh and eighth grade science, as well as psychology, in Swallow's combined secondary school. He has almost 30 years of experience as a middle school science teacher, coaches wrestling, and announces for home football games. He was formerly a football and track coach in the high school and at one time was a Catholic school principal. Paul looks to be of late middle age, with a wiry build and strong handshake. He was the first teacher interviewed in this study and was an engaged participant who quickly responded to emails asking for clarification. But he often conveyed skepticism about the usefulness of studying what different teachers did inside their classrooms. We met twice to discuss his classroom practice, and I observed him again during a K-12 meeting when all science teachers across the district's two schools had a round-table discussion of their MEAP results. He is not a member of MSTA, nor has he participated in the MEAP development process.

Like most teachers in this study, Paul seemed comfortable around students. While visiting with him I saw students leaving his room after class, and on one occasion while we were talking during his planning period, two boys in school sports jackets came in and sat down at one end of the room, seemingly to hang out. When I mentioned their presence, he said something along the lines of "hey you two get out" and they rose to leave. When I said they weren't a problem for me, he said, "hey" shrugged, and waved his hand at the kids, who went back to sit down at the edge of the classroom.

Classroom

Paul's classroom is neat and organized. The rectangular tables have hard black tops and room for two students. At the front of the room is a large desk, and along the back wall is a Periodic Table of the Elements (see Photo 5.1) that was made by students, with each element represented by yarn and beads on letter-sized paper. Under a television at the front of the room hang various neckties. Student work and projects are visible throughout the room. Particularly noticeable is a model of the solar system pasted on the ceiling (Photo 5.2). A central sun radiates in different strips of paper. At various points along each strip hang paper discs to indicate planets and signs to indicate relevant facts. In some places, dried beans and other items pasted on paper represent asteroid belts. This model helps to represent the scale of the solar system more accurately than conventional textbook drawings, which often fail to show size and distances correctly. Along the window wall is a shelf with hundreds of colored folders. These are the journals for the students. As with many teachers in this study, the journals of Paul's students usually stay in the classroom. On the same wall is a tall filing cabinet from which Paul retrieved tests and quizzes when I asked for examples.

Photo 5.1 - Wall of Paul Bond's classroom showing atom models made from yarn and beads

Photo 5.2 - A paper model of the solar system on Paul Bond's classroom ceiling

<u>Philosophy of Teaching and Assessments</u>

Paul believes that classroom approaches are individual to a teacher and that what he does should not be forced on other teachers. He said his approach to teaching science is different from many in the building. He said he is less fact-based and does not have kids do big projects at home because he has found the parents often do the projects. When asked how assessments might help him do what he cares most about doing in his job, he gave this response in Transcript 5.1.

Transcript 5.1 - Paul Bond discusses his goal for science education.	
1)	PB:That I care about the most in my job… The thing that I want most in my job is that kids get the big picture of science Its like when I plan a lesson, I say OK we gotta teach, for example seventh grade is on minerals now I don't look and say I want them to know about what Micah I mean Pyrite and nanannana I go into it looking at, OK we've got characteristics of minerals. What experiment can we do that will make this REAL to them. Y'know...I'm looking for OWNERSHIP of the ..I want them to have it in here (pounds heart) not just be able spit out facts You always have the kids that can spit out facts no matter what But somewhere down the road .. I don't expect them to remember oh the name of that mineral is…but they'll remember the Mohs hardness scale because we've tested all these minerals for hardness. We've done the street test we've looked at luster and all these things. And I think that to me that's more important than memorizing the facts is to have ownership of the concept.
2)	PP: And they will see those concepts in their assessments?
3)	PB: Exactly.

Paul does use the textbook, but does not rely on it exclusively. While he finds the textbook (Glencoe) well aligned with the state standards he teaches to, he often does not prefer the order of the lessons in the book. He also mentioned Montessori systems when asked for a new science assessment that would be beneficial. He thought that if school work could be more student-directed, as it may be in home school situations, and less directed by him, many students would benefit.

Assessment System

While Paul has developed most of his assessment approach himself, he takes parts from other documents, including his Glencoe textbook, and rearranges them to fit his own needs. He mentioned that he learned different techniques from other teachers at workshops and borrowed from other sources where needed. For example, a recent quiz on the solar system had 80% of the questions from the textbook and 20% from him. He said this mixture was typical. While the kids are usually engaged in some activity that is assessed, Paul gives them a few experiences without assessments, such as a model rocket activity. Table 5.1 describes the key elements of Paul’s assessment system.

Table 5.1 - Assessment instruments used in Paul Bond's classroom

Assessment activity	Activity Frequency	Duration	Activity Relationships	Activity Source	Comments
A. Bellwork* (Logon question)	Daily	1 min.	Into journal	Teacher created	Often topical.
B. Experiment Worksheet*	100 per year (every 1-2 days)	5 min.	Into journal	Teacher created	25% of grade. He says kids love these.
C. Quiz*	Weekly	5-15 min.	Preview of test. Into journal	Teacher created	25% of grade
D. Journal*	Daily	Few min.		Student	25% of grade
E. Unit tests	Approx 2 months	Class period	Into journal	Textbook Supplement	Combine for
F. Final Exams	End of Term	Class period plus prep time	Additional assessment	Teacher created	25% of grade

* - examples provided.

Paul's assessment system has the six elements shown in Table 5.1. They can be considered a system because Paul arranges them in specific, interrelated ways. Almost daily, the kids are given a short question that Paul calls the *logon question*, but that could more generally be called *bellwork*, as it is usually a task that occurs as the kids are coming into class and the bells may still be ringing. Next on this list, occurring every day or two, is another short task that Paul calls the *experiment worksheet.* This is Paul's own invention, in which the students fill out a template for an experiment to answer some question he has assigned. The experiment worksheet is an important part of his teaching. It is where students frame their understanding of a problem in terms of evidence and process. He provides them with a rubric for his grading of this key activity. He stated that by the end of the year, a student will have done 100 worksheets. These first two items, along with short quizzes, go into a journal that is graded once a year. In addition to unit tests that come entirely or in parts from a textbook supplement, Paul has a formal examination that includes a couple of days of prep.

Figure 5.1 provides a temporal description of Paul's assessments. This illustration, along with others for each of the seven teachers, is conceptual. The illustrations use a diagram approach that began as a reconstruction based on notes and interviews. All of the teachers were then given drafts of the illustrations for comment. All did provide some feedback. Consistent with his interviews, Paul's review seemed to indicate stability in his approach. He wrote, "Phil, this is exactly what I do," when he sent back the draft illustration for Figure 5.1 with minor changes. However, it must be noted that none of these illustrations is absolute. Many different types of events, such as weather, emergencies, and the variation in students from one year to another, often alter teachers' plans.

Document set 5.1 shows examples of Paul's assessment documents. The assignments frequently involve interpreting canonical visual representations and development by students of their own visual depictions of processes and entities that they study. Paul's assessments often include humor, such as a caricature of Einstein, and he referred to his classroom as the "Psych Ward" or the "Snake Pit," the latter because of his interest in reptiles. I was unable to get copies of Paul's tests or exams despite several requests. One item in this set is a special communication form that Paul sends home with students to have their parent or guardian sign. It is a way he developed to communicate with them about how their kids are performing in his class.

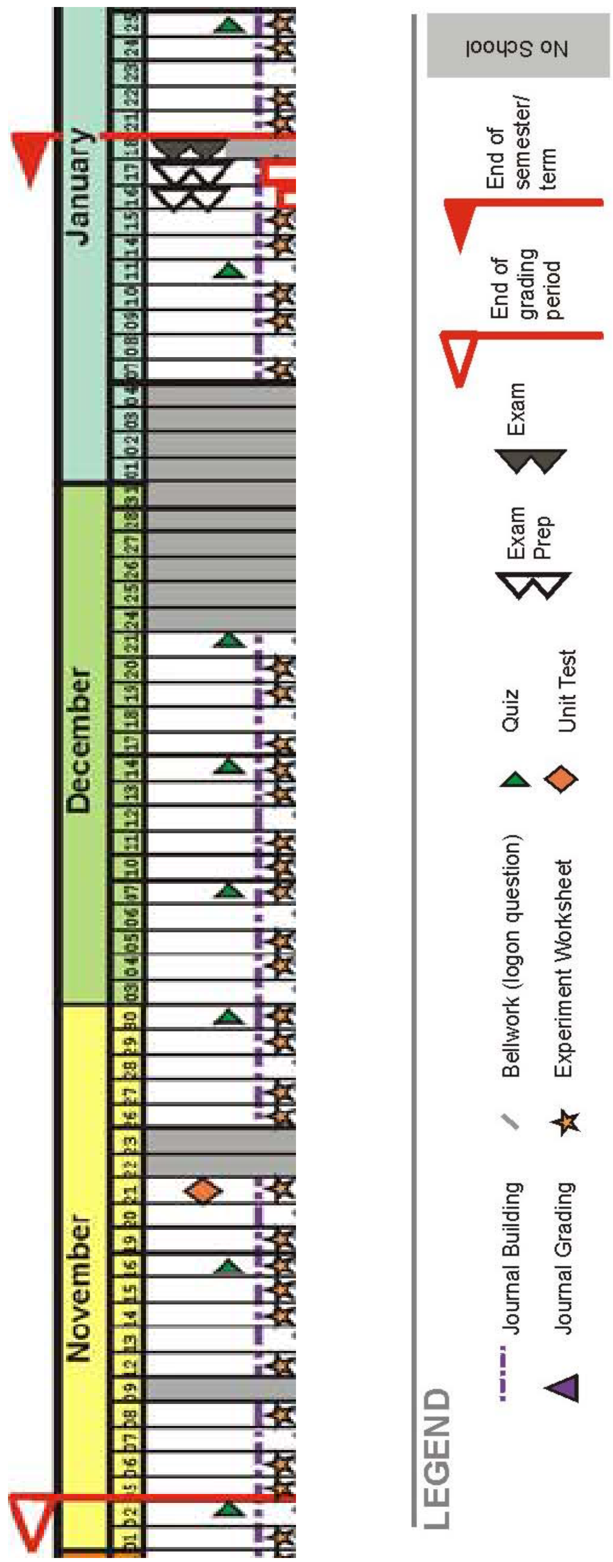
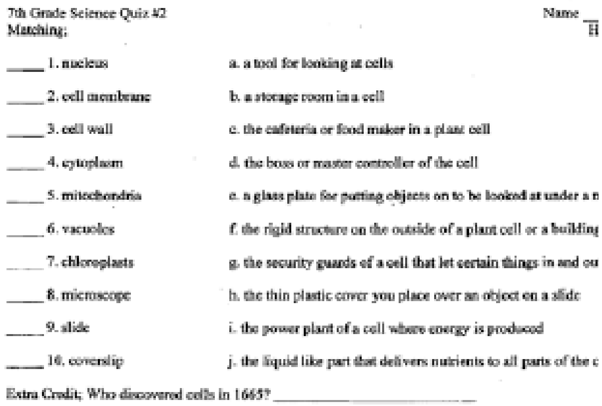


Figure 5.1 - Illustration of Paul Bond's assessment system over time.

a) Weekly quiz



7th Grade Science Quiz #2 Name ____

Matching:

____ 1. nucleus a. a tool for looking at cells

____ 2. cell membrane b. a storage room in a cell

____ 3. cell wall c. the cafeteria or food maker in a plant cell

____ 4. cytoplasm d. the boss or master controller of the cell

____ 5. mitochondria e. a glass plate for putting objects on to be looked at under a m

____ 6. vacuoles f. the rigid structure on the outside of a plant cell or a building

____ 7. chloroplasts g. the security guards of a cell that let certain things in and ou

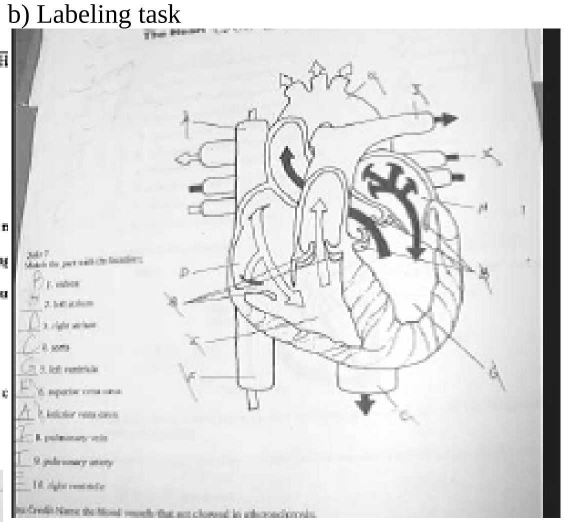
____ 8. microscope h. the thin plastic cover you place over an object on a slide

____ 9. slide i. the power plant of a cell where energy is produced

____ 10. coverslip j. the liquid like part that delivers nutrients to all parts of the c

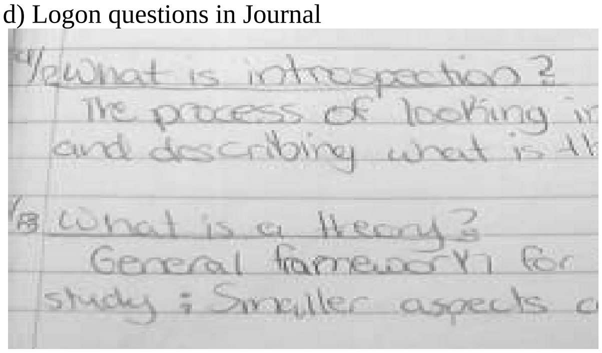
Extra Credit; Who discovered cells in 1665? ____________

b) Labeling task



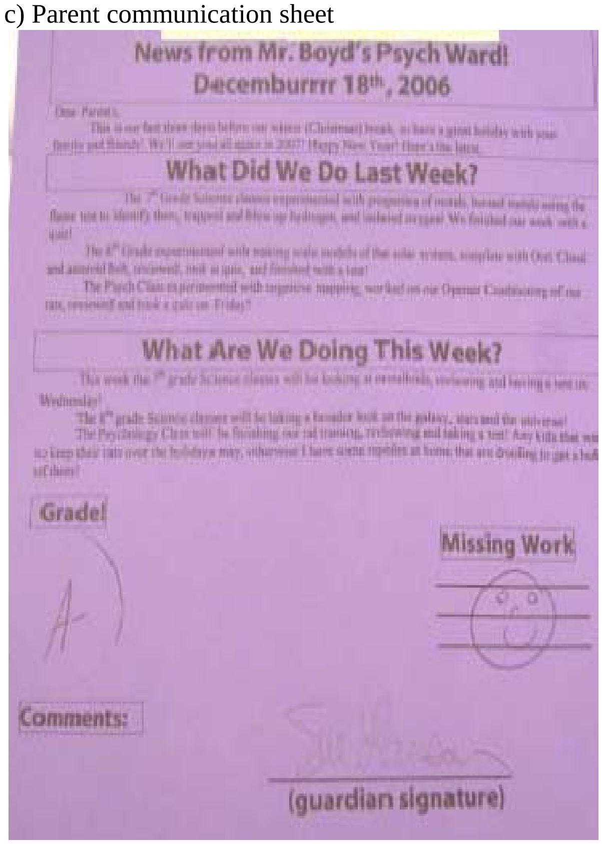
2. left atrium
3. right atrium
4. aorta
5. left ventricle
6. superior vena cava
7. inferior vena cava
8. pulmonary vein
9. pulmonary artery
10. right ventricle

d) Logon questions in Journal



What is introspection?
The process of looking in and describing what is th
What is a theory?
General framework for study ; Smaller aspects c

c) Parent communication sheet



News from Mr. Boyd's Psych Ward!
Decemburrrr 18th, 2006

Dear Parents,

What Did We Do Last Week?

What Are We Doing This Week?

Grade!

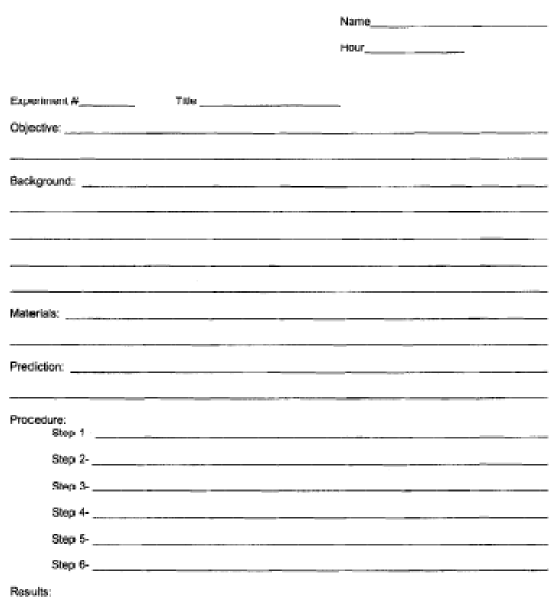
A-

Missing Work

Comments:

(guardian signature)

e) Experiment worksheet



Name____________

Hour____________

Experiment #________ Title ____________

Objective: ____________

Background: ____________

Materials: ____________

Prediction: ____________

Procedure:
Step 1 ____________
Step 2- ____________
Step 3- ____________
Step 4- ____________
Step 5- ____________
Step 6- ____________

Results:

Document set 5.1 - Assessment texts from Paul Bond's classroom

The Case of Faith Churchill in Dorchester Middle School

Faith Churchill (FC) is in her 30th year of teaching. She currently teaches seventh grade science in Dorchester Middle School. She began her career teaching physical education and she also teaches fitness, as well as being the school's union representative.

Faith is a slight woman with an electric energy and a speaking style full of personal conviction. She sees seventh grade as an important time for her students' development and her class as pivotal for their future choices. On several occasions in our interviews, she discussed students who were not typically high achievers and told several stories about what she did to help these students see themselves as more capable. She seemed keenly aware of the transition that the students are making into young adulthood and kept the students busy with assignments and activities. She also uses a number of point systems, including points for personal responsibility. She described how it is important for her that her students connect what they learn in the classroom with the outside world and that they develop skills for being adults in addition to performing well in her class.

Faith relies heavily on her textbook and used resources provided by the publisher (Prentice Hall) for some of her assessments. During the year I studied her, she began experimenting with new assessment technology: an answer sheet scanning machine called the GradeMaster. While professing a low level of technological literacy, she appeared to have no difficulty incorporating this technology into her teaching. The only impact on her practice she reported was that she had to rewrite/retype the questions so that the students could use the pre-printed scanning sheets. Her classes also develop presentations using PowerPoint, and she was the only teacher in this study that put up a calendar on a personal web page that I could access.

Classroom

Faith's classroom has two-seat black-top tables aligned in rows, with lab sinks and cabinets that line the perimeter. In addition to the "No Gum" sign over the entrance to the room, there are small signs that say "CHARACTER is who you are when no one is watching" and "SUCCESS IS A CHOICE,“ on the front of a large wooden desk/bench at the head of the room. On a cabinet that students would pass on the way into the classroom are five letter-sized bins labeled Monday through Friday that hold daily homework assignments.   While no student work is displayed in the room, Faith described how she used the cafeteria for one of her regular projects, where students build working model hot air balloons. She would mark how high the student project balloons rise on the wall itself. She frequently takes the students outdoors to a nearby stream or in the woods to observe local flora/fauna.

Photo 5.3 - Faith Churchill in the front of her classroom

Photo 5.4 - The daily homework bins in Faith Churchill's classroom and markers for students

Teaching and Assessments Philosophy

Faith's approach to teaching is to have the students work hard and hold them accountable for their performance. Rather than correcting their homework, she has the students correct their own so they have the additional practice of correcting it and she does not have to correct it herself. When the students enter the classroom, the answers to the homework are on the board, and they pass around a box of markers to use. If they do not correct it properly, she deducts points. In Transcript 5.2, she explained her rationale for this homework approach, which is like bellwork.

Transcript 5.2 - Faith Churchill describes the homework boxes in her classroom (from an email)

When students are absent they need not come to me and ask what we covered on that particular day they go right to the homework boxes and pick up their work. Of course if they are gone more than a week they know to look under the box for the previous weeks work. This also helps me just go and grab work if parents call to have homework sent home for particular days that their student has been absent.

This is also a great tool for students who forget their work at home--- they come in first thing in the morning, get the homework and take it to advisory or sit in my room and redo their work.

Basic concept—student accountability and responsibility.

She told many stories in our interviews about her approach to teaching. She seemed to have a soft spot for students who had difficulty and she was interested in providing opportunities for the students to see themselves as capable and able to achieve. She described this principle in the story in Transcript 5.3.

Transcript 5.3 - Faith Churchill discusses a student who needs hands on activity.

1) FC: think the engineering part is just because you have kids who are just aren't real good in the book but they can put together something so quick.
I had a kid last .. a couple of years ago. We used to fight all the time, he was a problem child.
But we were but he still comes down to see me all the time.
We finally .. got to see things each other's way –my way.
2) PP: Nice to have those compromises.
3) FC: He just was not he is not a book kid. He is all about tactile. He works with his grandpa. His grandpa owns an um um a gravel pit and he works with mechanical things. And I was doing a speed .. velocity you know all these different things and I had these cars and two of them I could just not get em to work and I'm pretty mechanical.

Transcript 5.3 - Faith Churchill discusses a student who needs hands on activity.

And so I looked at him and went .. here can you fix that ? (snap's fingers) 20 seconds.

4) PP: Really

5) FC: He said do you have a paper clip? I said yeah. No that's too big I need a smaller one. Here. (snap) done.
I said you're so good at that big shot try this one. (snap) done.

6) PP: Wow.

7) FC: And I had torn it apart and put it back together again and still couldn't get it. And that's the kind of kid when you do these kinds of activities, they just shine.

8) PP: So it's important to them .. uh to have those kinds of activities .. not necessarily for all kids.. all kids may enjoy them but there are going to be some kids.

9) FC: They need that.

Faith thinks about the role assessments play in helping students make connections between what occurs in and outside of school, and she described it in Transcript 5.4.

Transcript 5.4 - Faith Churchill discusses students making connections through her assessments

1) PP: What about the uses and purposes for assessments outside the classroom

2) FC: Well, I think the kids um I think it just helps y'know. I am gonna relate to the MEAP again if they score well, the parents and the community outside .. they buy into the school more .. and they don't think you're the bad mean teacher um if your MEAP scores are high y'know all the grief they've talked about and the kids it's a good thing.

3) PP: Well what about the tests they take in here, the quizzes the things like that? Are those mainly for your use in here or do you see uses for them outside?

4) FC: Well there are a lot of them just for in here .. but I will tell you .. y'know we study the phases of the moon and the kids will come in and say *did you see **that moon**? It was in a waxing crescent the other night it was gorgeous.*
y'know that's enough .. for me .. that's enough for me y'know
Or the sun *did you see the **sun** this morning* I mean the other day it was just this **huge** ball right on the horizon.' and the kids said *did you see the **photosphere**? The photosphere was just **blaring** had to put my glasses on.* That's enough for me. Y'know when they start looking and identifying.

5) PP: And making that connection?

6) FC: That connection, yup, and that's where my assessments have done um making sure they get that making sure they are aware of it.

<u>Assessment System</u>

Faith's assessment system, like Paul Bond's, features regular tests and quizzes. The test occurs approximately every two weeks and quizzes are weekly. The schedule is somewhat flexible, and she decides to have a test when she feels the class has covered the "right amount of material." In some cases she stretches a unit to avoid beginning a new unit just before a break. She does not have a formal exam and instead uses the non-cumulative test as the focal assessment. The tests often come from resources provided by the textbook publisher, while she develops the quizzes. She said she is currently revising some assessments to use the GradeMaster scanner.

Table 5.2 - Assessment activities for Faith Churchill

Assessment activity	Activity Frequency	Duration	Activity Relationships	Activity Source	Comments
Homework	Almost daily	Few min.		Teacher	Students correct it.
Quiz	2 weeks. Usually between tests	10 min.	Prefigures the test	Teacher	
Practice Quiz	Usually 2-3 days into the unit	Few min.	Acts similar to a pretest	Teacher	Sometime a special area such as math
Test	Usually every 2 weeks	Up to one class	Caps unit	Prentice Hall	
Lab/Project	Once per unit. (3 units have big ones)	Varies	Integrates lesson material	Teacher defines, students build	Usually done in teams
Journals	Daily	Daily	Includes other work	Students build	

Like Paul, Faith arranges her instruments into a predictable system. By contrast, she has more frequent tests and regular quizzes including practice (informal) quizzes. She assigns homework almost daily, whereas Paul rarely assigns it. She has a lab or project for each unit, something that is not part of Paul's practice. Some projects involve going outside. Others are indoors. Figure 5.2 is a conceptual illustration of Faith's assessment system as it might occur over time.

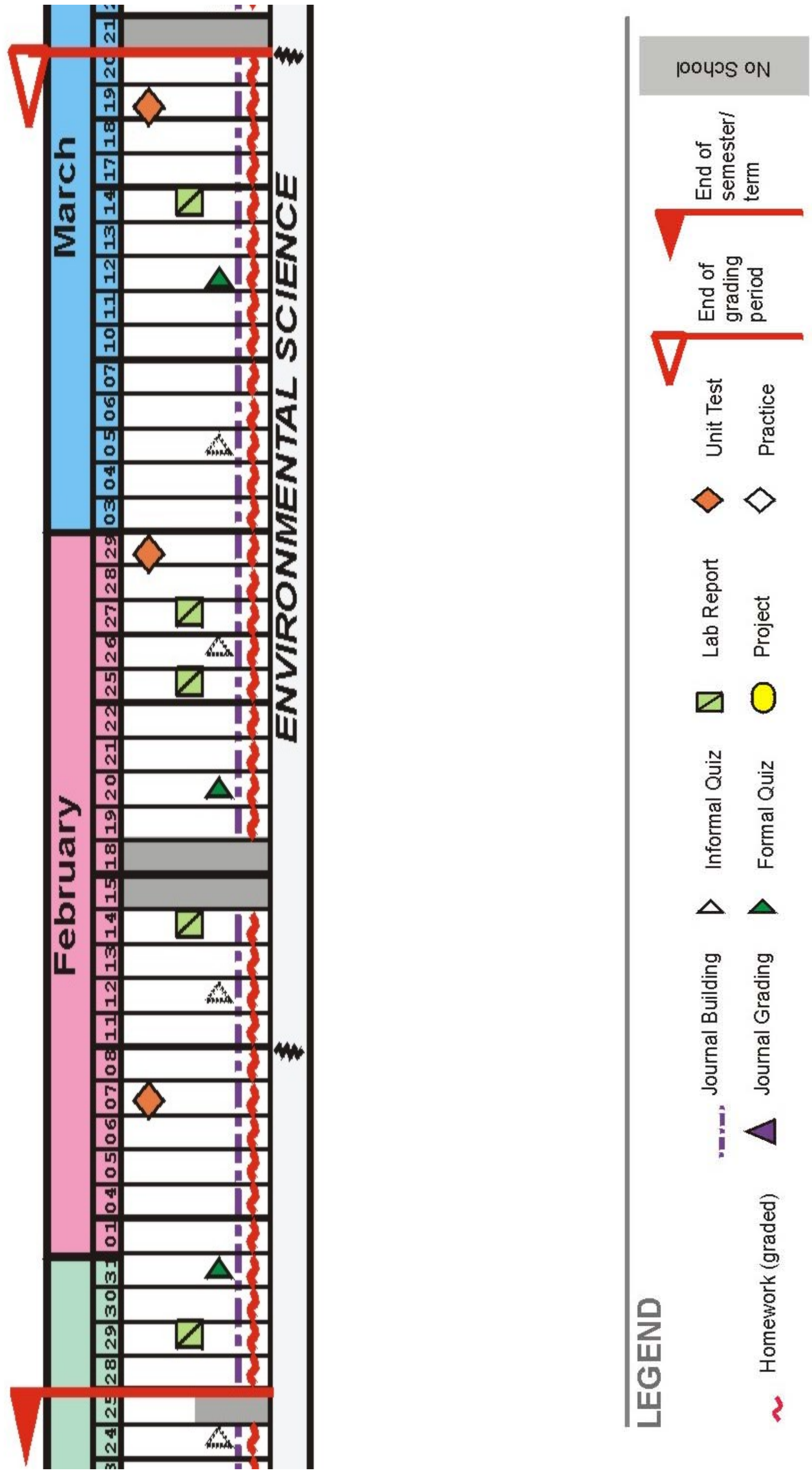
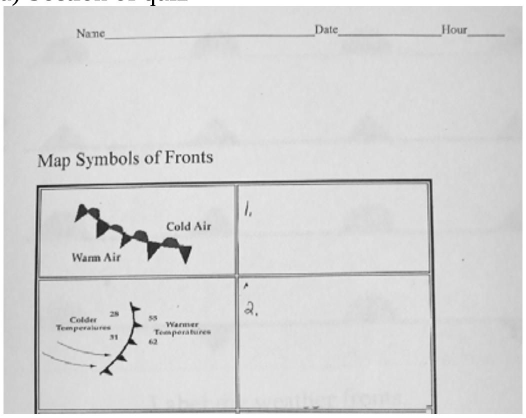


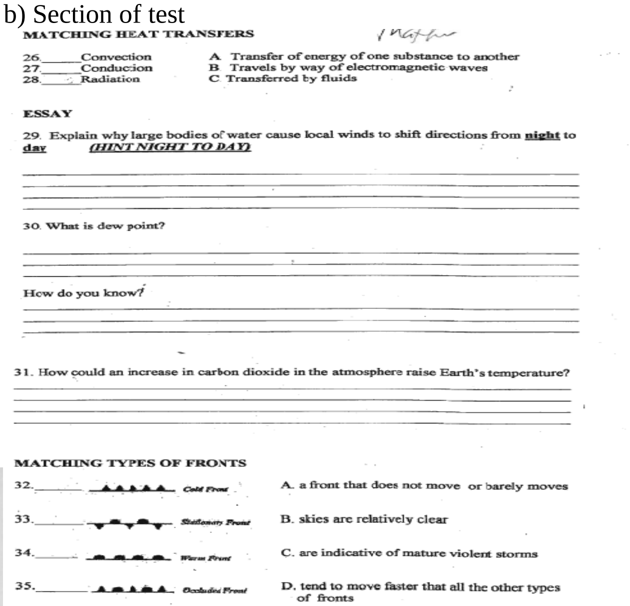
Figure 5.2 - Illustration of Faith Churchill’s assessment system over time

a) Section of quiz

b) Section of test


Name__________________________Date______________Hour_____

Map Symbols of Fronts


MATCHING HEAT TRANSFERS

26._____Convection A Transfer of energy of one substance to another
27._____Conduction B Travels by way of electromagnetic waves
28._____Radiation C Transferred by fluids

ESSAY

29. Explain why large bodies of water cause local winds to shift directions from **night** to **day** (HINT NIGHT TO DAY)

30. What is dew point?

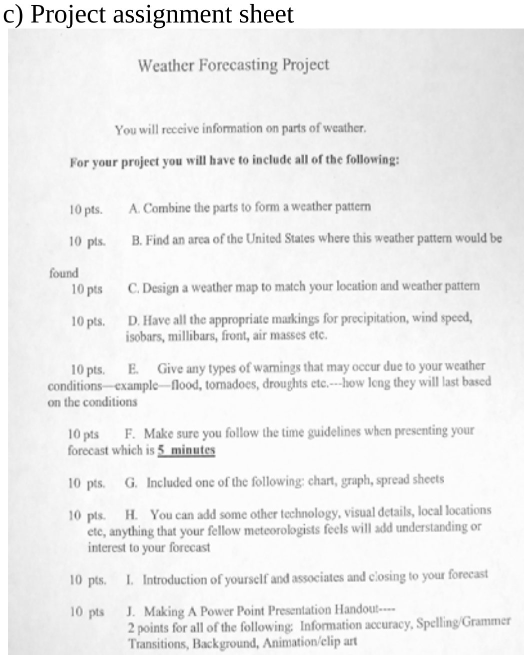
How do you know?

31. How could an increase in carbon dioxide in the atmosphere raise Earth's temperature?

MATCHING TYPES OF FRONTS

32._____ Cold Front A. a front that does not move or barely moves
33._____ Stationary Front B. skies are relatively clear
34._____ Warm Front C. are indicative of mature violent storms
35._____ Occluded Front D. tend to move faster that all the other types of fronts

c) Project assignment sheet


Weather Forecasting Project

You will receive information on parts of weather.

For your project you will have to include all of the following:

10 pts. A. Combine the parts to form a weather pattern

10 pts. B. Find an area of the United States where this weather pattern would be found

10 pts C. Design a weather map to match your location and weather pattern

10 pts. D. Have all the appropriate markings for precipitation, wind speed, isobars, millibars, front, air masses etc.

10 pts. E. Give any types of warnings that may occur due to your weather conditions—example—flood, tornadoes, droughts etc.---how long they will last based on the conditions

10 pts F. Make sure you follow the time guidelines when presenting your forecast which is 5 minutes

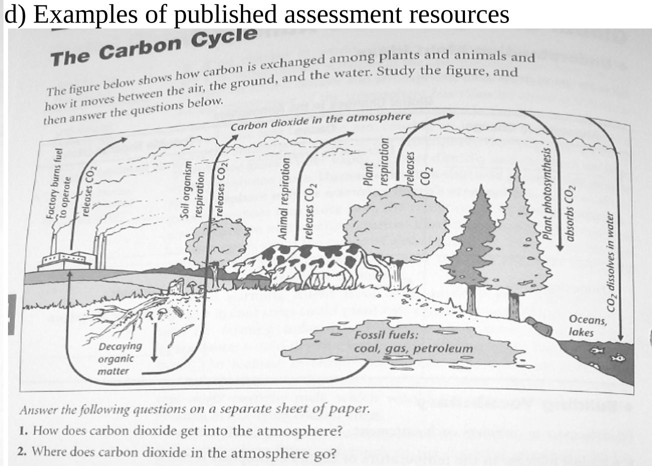
10 pts. G. Included one of the following: chart, graph, spread sheets

10 pts. H. You can add some other technology, visual details, local locations etc, anything that your fellow meteorologists feels will add understanding or interest to your forecast

10 pts. I. Introduction of yourself and associates and closing to your forecast

10 pts J. Making A Power Point Presentation Handout----
2 points for all of the following: Information accuracy, Spelling/Grammer Transitions, Background, Animation/clip art

d) Examples of published assessment resources


The Carbon Cycle

The figure below shows how carbon is exchanged among plants and animals and how it moves between the air, the ground, and the water. Study the figure, and then answer the questions below.

Answer the following questions on a separate sheet of paper.

1. How does carbon dioxide get into the atmosphere?
2. Where does carbon dioxide in the atmosphere go?

Document set 5.2 - Assessment texts from Faith Churchill's classroom

The Case of Betsy Dearing, Avon Falls Middle School

Although not as experienced as Paul or Faith, Betsy Dearing (BD) presents an example of a teacher very involved with her students and her school. Betsy teaches seventh grade science in Avon Falls Middle School and when we met, she was in her seventh year of teaching. She is currently the school's department head and has been active in and presented at the MSTA. She also leads school trips, including an outdoor weekend camp and a space camp field trip to Alabama. She was a supportive study participant and met with me on several occasions. Once she stayed late past 6:00 p.m. to answer questions, when her school day had ended at 3:30. She worked in outdoor and environmental education before becoming a school teacher.

Energetic and attentive, Betsy seems to be in her late twenties or early thirties. She was engaged in and reflective about the study, in that she brought examples of assessments with notes she had made on them to help me with my research. She discussed her revision process and how she uses various resources in making up her assignments and assessments. Betsy also teaches using a Prentice Hall textbook, and each unit she teaches is organized around a book chapter. Her curriculum is largely focused on life science, with an emphasis on the body and nutrition, as well as a short unit on sound and light. Like other teachers in this study, she seemed comfortable around students.

Photo 5.5 - Counter and cabinets in Betsy Dearing's classroom with classroom pets

Classroom

In Betsy’s classroom the student desks are individual. The room has wooden cabinets with black lab counters with sinks along the perimeter, and almost all of the cabinets are covered with posters. Some of the counters hold aquariums and terrariums with various classroom pets that were donated by parents of students. Along the top of one long wall in the back of the room are long strips of cash register paper that have been colored. These are student models of the intestines shown in their actual length. Like Paul Bond, she uses the building as a textual space to illustrate various concepts of scale.

Photo 5.6 - Intestine models made from register paper in Betsy Dearing's classroom

Teaching and Assessment Philosophy

Betsy did not explicitly state a philosophy of teaching or assessments in the way that Paul or Faith did. But, from a number of her responses and my study of her documents, I am able to identify several themes. First, she tries to make connections between her class and students' home lives. She asks kids to look at the things they eat at home and analyze them for nutrition and fat content. She also designed assignments to relate to popular media. One was called the "Top Chef" assignment (after the title of a popular TV show) and integrated work the students had been doing in class with homework. She placed a lot of emphasis on student development, both in their body and in the emotional changes that seventh graders can be going through. In other discussions about the design of assessments, she said it is important to her to give students multiple opportunities to succeed. Like Faith, she said they needed to be responsible and have consequences for lack of responsibility.

Assessment System

Betsy's assessment system, summarized in Table 5.3, is something she invested time in developing and changing. She shared with me her revision process and rationale for assigning certain content to particular instruments. As she designed assessments, she was attentive to how much students can learn or how they may be confused by similar but different concepts. In one case, she used an extra assessment to reinforce the separation of the concepts of habitat and food. For example, in a habitat, food is available to eat; and living things in habitats need food to eat. These are similar and related concepts that the textbook puts into adjacent units. For clarification, she used separate quizzes to make sure kids understood the differences.

She created her own quizzes and tests; often using material from textbooks, the Internet, and any other sources she finds. One area she discussed having a constant need for help with is appropriate graphics. Her quizzes tended to preview the material and visual representations that the test covered so that the quiz operated similar to a test practice. When she did assign homework, it could be project-like, such as when she asked students to do research in their home environments to be used later in class. Just before the end of the term, she collects and grades the journals. She said the journal can be a “grade maker or a grade breaker” depending on the effort and organization it shows. While Betsy did not describe any assessments as projects, the nature of her homework assignments would tend toward projects, as they occur over several classes and involve data.

Table 5.3 - Assessment activities for Betsy Dearing

Assessment Activity	Activity Frequency	Duration	Activity Relationships	Activity Source	Comments
A. Homework /assignments	Continuous throughout the semester	1/3 of class time	Integrated activity organized by unit.	Textbook and Betsy	These can involve projects that include data collection
B. Quiz (formal and informal)	Usually one per unit before test	15 min.	Previews some of the test (a practice)	Betsy	Not all units have one
C. Test	Approx. monthly. One per unit	A class period	Quiz is a practice	Betsy	
D. Extra quiz	Occasionally	A class period (usually student time)	Used to separate content for students	Betsy	When terms & concepts are duplicated
E. Journals	Yearly. Checked once during marking period		Includes homework and assignments	Students maintain them	Used for notes and vocabulary.

* examples provided.

Figure 5.3 is a representation of how Betsy’s assessment system might look like from a temporal perspective. The schedule for her assessments/assignments fluctuated based on the year and what was occurring at different points in the school calendar. In these ways, her approach was different from Paul and Faith, who tended to have equal units and regular pacing intervals. (This aspect will be discussed later in this chapter when the units that teachers covered are represented in Figure 5.8). She spends about two to three weeks on the shortest (heredity) and twelve weeks on the longest (human body). Her school year began with a review of measurement-oriented tasks, and she said she tried to cover most of the hard material before spring break. Similar to Faith, she tried not to hold untested material over a break and may create an assessment to give just before the students leave on break.

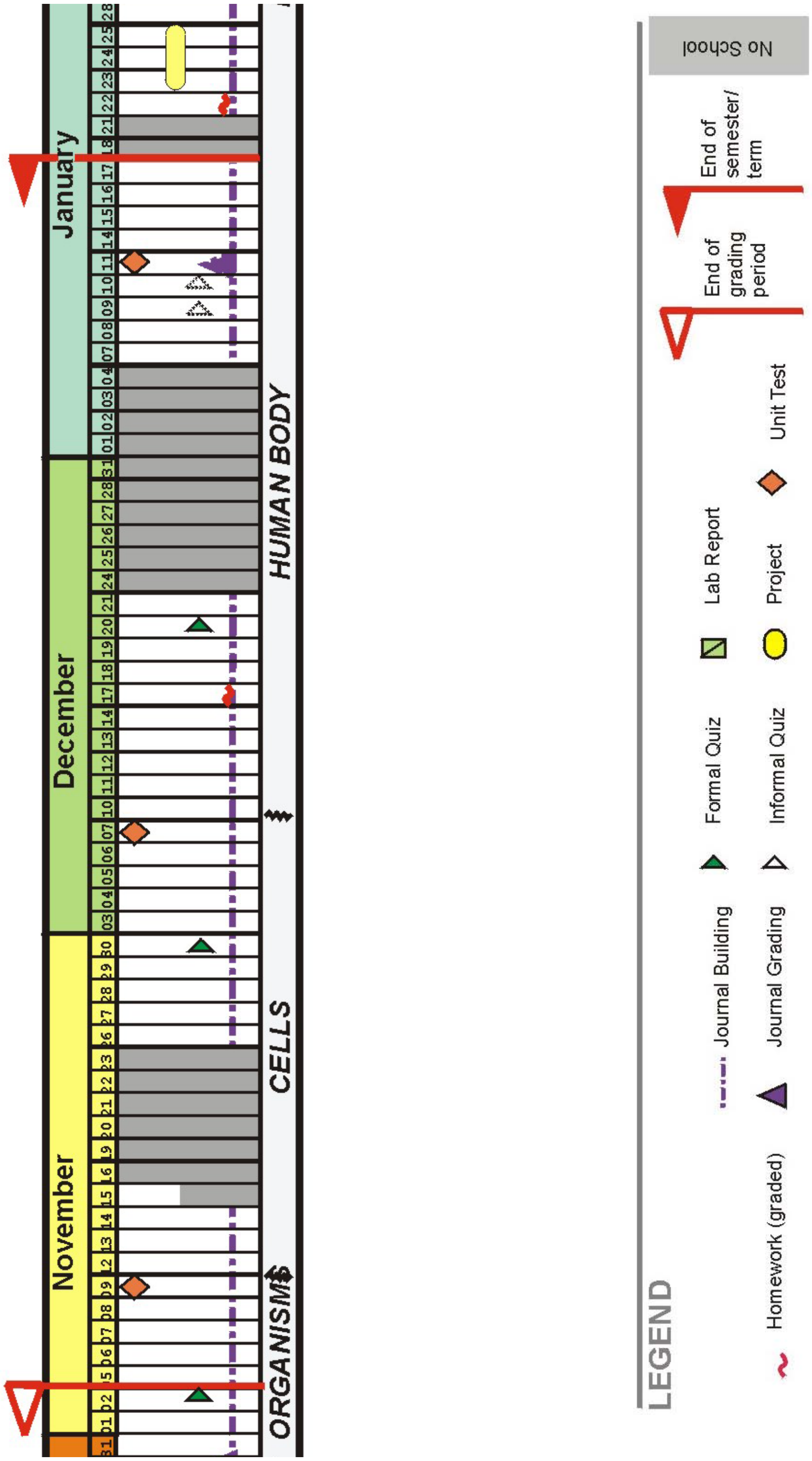
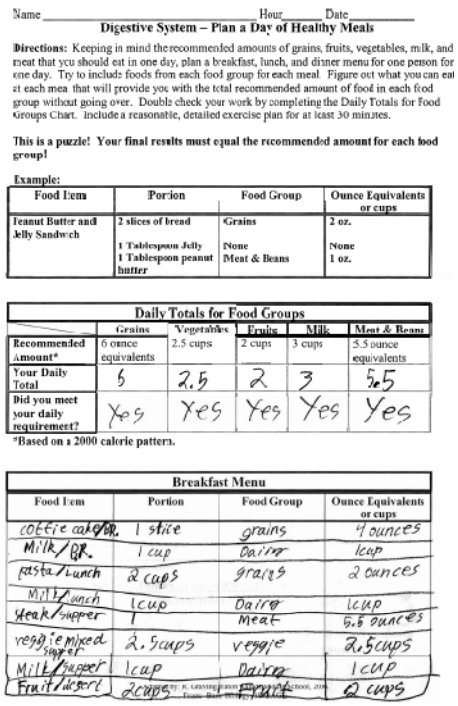


Figure 5.3 - Illustration of Betsy Dearing's assessment system over time

a) Homework/project worksheet


Name ______________________ Hour______ Date__________

Digestive System – Plan a Day of Healthy Meals

Directions: Keeping in mind the recommended amounts of grains, fruits, vegetables, milk, and meat that you should eat in one day, plan a breakfast, lunch, and dinner menu for one person for one day. Try to include foods from each food group for each meal. Figure out what you can eat at each meal that will provide you with the total recommended amount of food in each food group without going over. Double check your work by completing the Daily Totals for Food Groups Chart. Include a reasonable, detailed exercise plan for at least 30 minutes.

This is a puzzle! Your final results must equal the recommended amount for each food group!

Example:

Food Item	Portion	Food Group	Ounce Equivalents or cups
Peanut Butter and Jelly Sandwich	2 slices of bread 1 Tablespoon Jelly 1 Tablespoon peanut butter	Grains None Meat & Beans	2 oz. None 1 oz.

Daily Totals for Food Groups

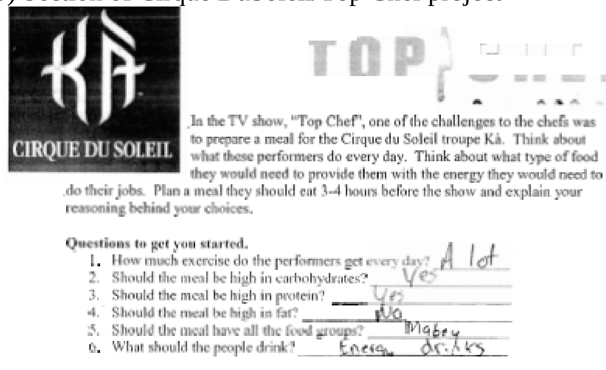
	Grains	Vegetables	Fruits	Milk	Meat & Beans
Recommended Amount*	6 ounce equivalents	2.5 cups	2 cups	3 cups	5.5 ounce equivalents
Your Daily Total	6	2.5	2	3	5.5
Did you meet your daily requirement?	Yes	Yes	Yes	Yes	Yes

*Based on a 2000 calorie pattern.

Breakfast Menu

Food Item	Portion	Food Group	Ounce Equivalents or cups
coffie cake/BR.	1 slice	grains	4 ounces
Milk/BR.	1 cup	Dairy	1cup
pasta/Lunch	2 cups	grains	2 ounces
Milk/Lunch	1cup	Dairy	1cup
steak/supper	1	Meat	5.5 ounces
veggie mixed supper	2.5cups	veggie	2.5cups
Milk/supper	1cup	Dairy	1cup
Fruit/desert	2cups	Fruit	2 cups

b) Section of Cirque DuSoleil/Top Chef project

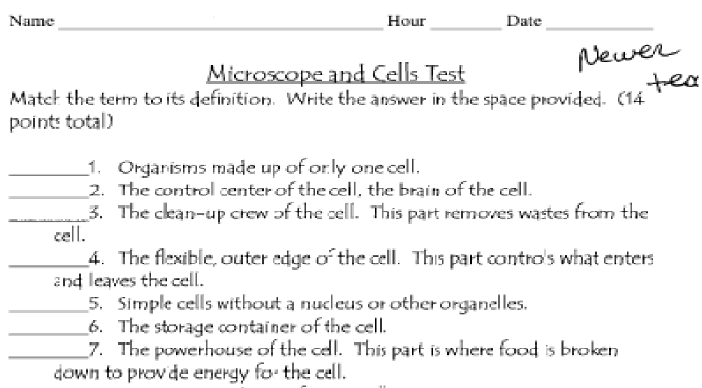


In the TV show, "Top Chef", one of the challenges to the chefs was to prepare a meal for the Cirque du Soleil troupe Kà. Think about what these performers do every day. Think about what type of food they would need to provide them with the energy they would need to do their jobs. Plan a meal they should eat 3-4 hours before the show and explain your reasoning behind your choices.

Questions to get you started.

1. How much exercise do the performers get every day? A lot
2. Should the meal be high in carbohydrates? Yes
3. Should the meal be high in protein? Yes
4. Should the meal be high in fat? No
5. Should the meal have all the food groups? Mabey
6. What should the people drink? Energy drinks

d) Section of quiz


Name ______________________ Hour ________ Date __________

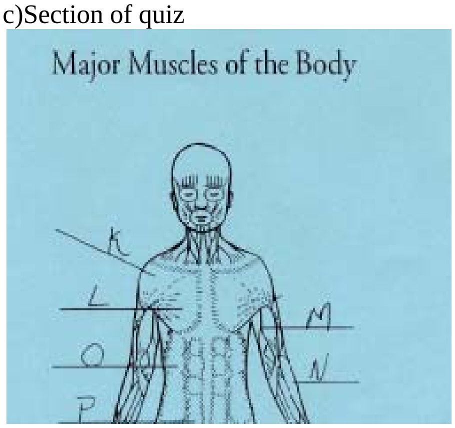
Newer test

Microscope and Cells Test

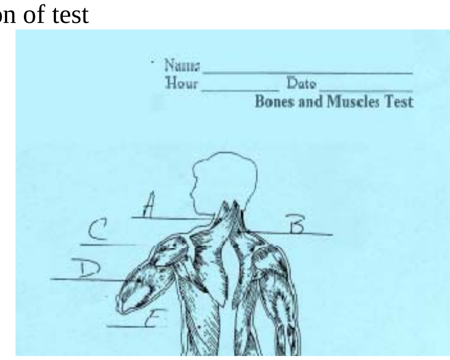
Match the term to its definition. Write the answer in the space provided. (14 points total)

_________1. Organisms made up of only one cell.
_________2. The control center of the cell, the brain of the cell.
_________3. The clean-up crew of the cell. This part removes wastes from the cell.
_________4. The flexible, outer edge of the cell. This part controls what enters and leaves the cell.
_________5. Simple cells without a nucleus or other organelles.
_________6. The storage container of the cell.
_________7. The powerhouse of the cell. This part is where food is broken down to provide energy for the cell.

c)Section of quiz



e)Section of test



Document set 5.3 - Assessment texts from Betsy Dearing's classroom

The Case of Jim Heinrich, Avon Falls Middle School

Jim Heinrich (JH) teaches seventh grade science next to Betsy Dearing in Avon Falls Middle School. He has been a public school teacher in the U.S. for seven years. Previously, he taught in private schools in Italy and Greece. He is not an MSTA member and was initially a hesitant participant in the study, but became more engaged as time went on.

Jim is a tall, late-middle-aged man with blond hair and a gentle demeanor. He was the only teacher who did not give permission to be taped, so these records rely more on field notes and recollection. He did, however, give many extended responses and often provided perspectives that others did not. As the study progressed, he was the only participant who inquired about me personally. Before we discussed his classroom work at the start of the interview process, he asked me about my research in detail and wanted to understand the motivation and purposes for it. Once that was discussed, he became increasingly engaged and responsive; asking questions and initiating discussions about fundamental issues such as the role of vocabulary in reading science texts or the underlying basis of dyslexia (both topics that are the focus of active educational research). He was also an active participant in the follow-up member checking via email in late 2007 through early 2008.

Classroom

Jim's classroom is long, with sets of black tables where students sit divided by a center aisle. The room is long and has wooden lab cabinets on both sides. On the way into Jim's room there is a display cabinet with various stuffed animals and examples of nature. The walls and cabinet doors are decorated with student work, including three-dimensional project models and posters. The room allows Jim to have a desk that is off to one side with a window view.

Photo 5.7 - Jim's desk near window

Jim is interested in music. He plays classical guitar and his room has large stereo speakers on the cabinets. On one occasion I met with him as he was finishing what seemed to be a music lesson or a discussion about music with a student. Jim said most concepts are best demonstrated with pictures. His whiteboard features drawings. He gives students assignments that require them to draw scientific representations. Like Betsy, he also has a unit on sound and light in addition to life science.

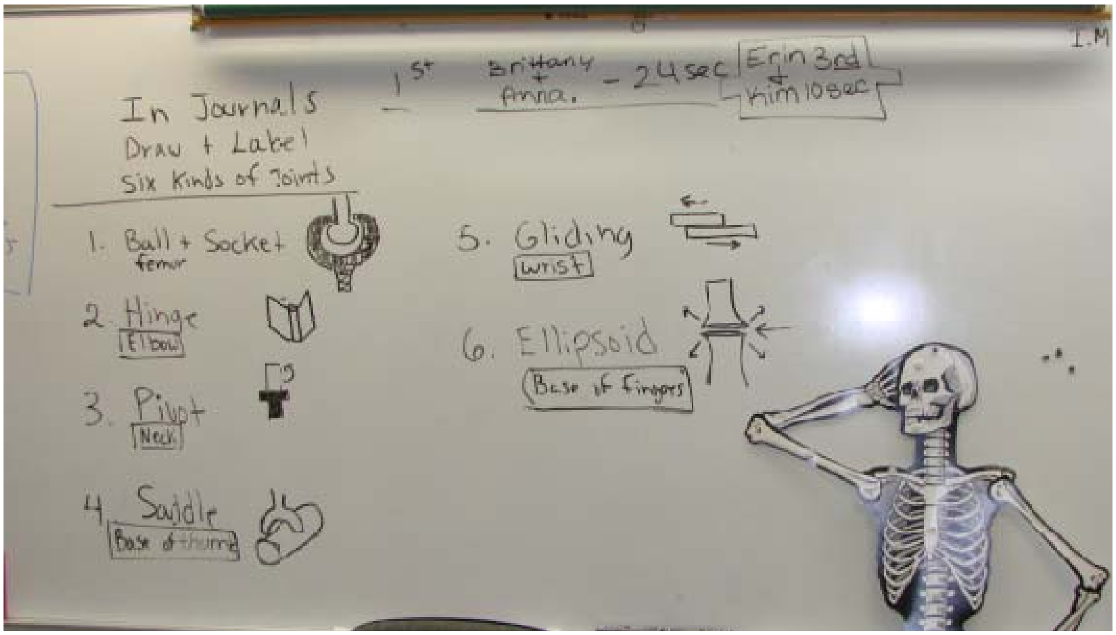


Photo 5.8 - Drawings and instructions on Jim's whiteboard

Teaching and Assessments Philosophy

Jim's teaching emphasizes factual content and does not include much essay work. In part, he said, this is because of time constraints. Jim believes in the importance of teaching fundamental concepts, including scientific terminology, in order for students to understand complex processes and said, "It is a building process; [students need to] build a foundation to be able to understand simple concepts to be able to eventually put together more complicated ideas." When discussing students, he described their propensity to forget and their need to have reinforcement to adequately build their conceptual structures. To provide this foundation, he assesses frequently and states that without assessments, students are likely to "have the knowledge leak out." In his teaching and assessments he uses what he calls *reword questions,* where students are asked to take a concept and state it in their own words. He also described how his assessments help him evaluate his own work. They are like “a scorecard” for the concepts he has presented. He said he sees himself like a coach and that the assessments gave him wins and losses in a contest for the students’ comprehension. He also expressed the belief that assessments were limited in what they could show about students’ underlying conceptual structures.

Assessment System

Jim’s classroom assessment was built around a weekly test and quiz. He sent home a practice sheet for students to prepare for the test and used quizzes and used oral feedback as a less formal assessment. Unlike Betsy, his assessment approach was regular. The practice was almost always given the day before the test, and a quiz (that also previews the test material) occurred earlier in the week. Jim regularly assigns homework. Like Betsy, he had students develop a journal. But I was unable to get any information on it being collected or graded. About six times a year he allows students to do an extra credit project, which usually takes the form of a poster. These projects are ways that students who need to raise their grades can compensate for sub-optimal performance during the unit. Some of the projects are on display in his classroom (see Photo 5.7 above).

Table 5.4 - Assessment activities for Jim Heinrich

Assessment Activity	Activity Frequency	Duration	Activity Relationships	Activity Source	Comments
A. Test*	Weekly	45 min.	Practice sheet	Created with textbook diagrams	
B. Test practice sheet*	Weekly	45 min.	Content same as test	Jim	Students use book and notes
C. Quiz*	Ad hoc	30 min.		Jim	At least one quiz is shared by Betsy & Jim.
D. Oral feedback	Ad hoc	15 min.			
E. Extra credit projects	Ad hoc		As many as six, one per unit.		Gives students chances to raise their grade
F. Journal	Daily	20-30 min.	Contains notes, assignments, and homework	Students maintain it	Graded once per semester

* examples provided.

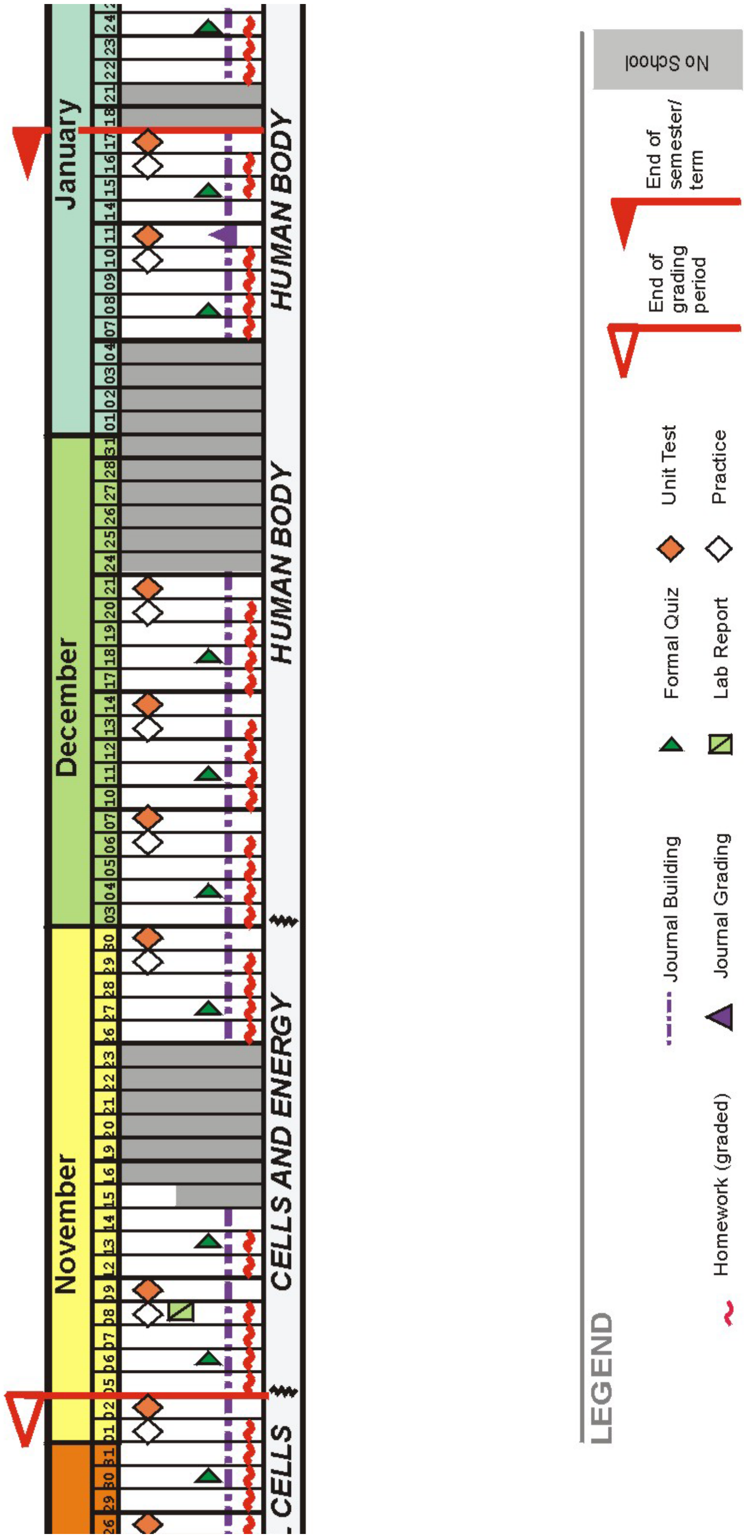
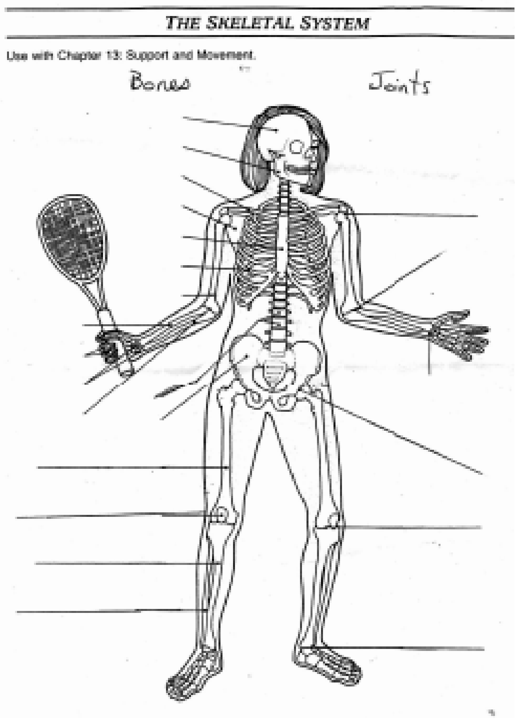


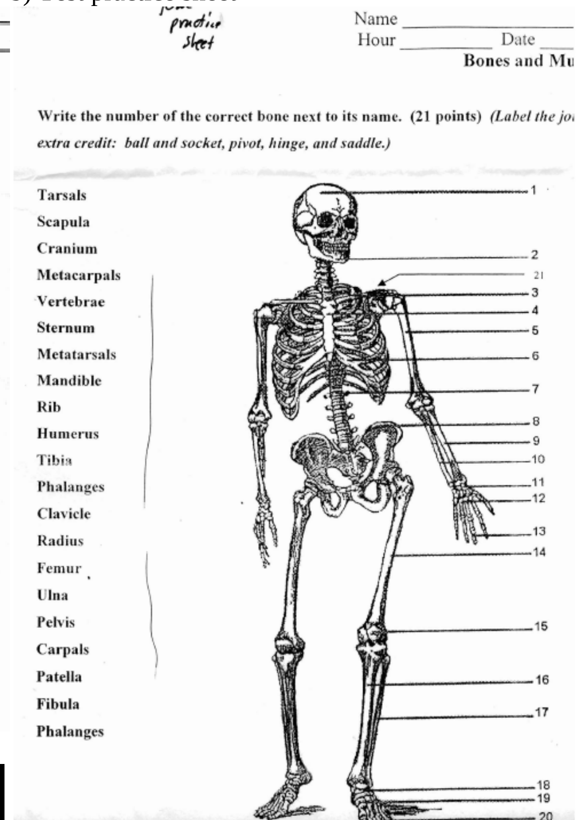
Figure 5.4 - Illustration of Jim Heinrich's assessment system over time

When displayed in temporal terms, Jim's assessment system has a regular, predicable rhythm, as Figure 5.4 shows.

a) Test



b) Test practice sheet

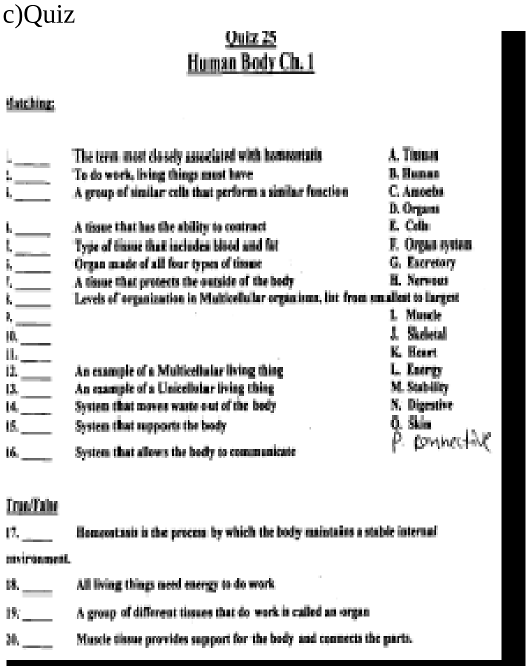

practice sheet

Name ______
Hour ______ Date ___
Bones and Mu

Write the number of the correct bone next to its name. (21 points) *(Label the jo extra credit: ball and socket, pivot, hinge, and saddle.)*

Tarsals
Scapula
Cranium
Metacarpals
Vertebrae
Sternum
Metatarsals
Mandible
Rib
Humerus
Tibia
Phalanges
Clavicle
Radius
Femur
Ulna
Pelvis
Carpals
Patella
Fibula
Phalanges

c)Quiz


Quiz 25
Human Body Ch. 1

Matching:

1. ____ The term most closely associated with homeostasis — A. Tissue
2. ____ To do work, living things must have — B. Human
3. ____ A group of similar cells that perform a similar function — C. Amoeba
D. Organs
4. ____ A tissue that has the ability to contract — E. Cells
5. ____ Type of tissue that includes blood and fat — F. Organ system
6. ____ Organ made of all four types of tissue — G. Excretory
7. ____ A tissue that protects the outside of the body — H. Nervous
8. ____ Levels of organization in Multicellular organisms, list from smallest to largest
9. ____ — I. Muscle
10. ____ — J. Skeletal
11. ____ — K. Heart
12. ____ An example of a Multicellular living thing — L. Energy
13. ____ An example of a Unicellular living thing — M. Stability
14. ____ System that moves waste out of the body — N. Digestive
15. ____ System that supports the body — O. Skin
P. connective
16. ____ System that allows the body to communicate

True/False

17. ____ Homeostasis is the process by which the body maintains a stable internal environment.
18. ____ All living things need energy to do work
19. ____ A group of different tissues that do work is called an organ
20. ____ Muscle tissue provides support for the body and connects the parts.

b) Student project

Document set 5.4 - Assessment texts from Jim Heinrich’s classroom

The Case of Christy Connolly, Crimson Middle School

There are two schools where I have two seventh grade science teachers as the basis for cases in this study. In addition to Betsy and Jim at Avon Falls, two teachers in Crimson Middle School, Christy Connolly (CC) and Ben Raminskis (BR), also participated in this study. Christy teaches both seventh grade science and mathematics. She was an enthusiastic participant. Christy is a short, middle-aged woman with seemingly constant energy. At the time of my study she was in her 12th year as a teacher. She left a career as an accountant and financial analyst. She is recognized in her district as an excellent science teacher and was selected Teacher of the Year by a state science professional association. She is a big fan of technology in the classroom and has been using the Moodle online course system[18] as a way to provide students with access to science content, assessments, and grades. Moodle is an open-source classroom technology similar to the software produced by Blackboard Corporation.[19] The district Crimson Middle School is within had a Moodle initiative for three years; it has been adopted unevenly by teachers. Christy was active in the MSTA at the time of this study and had served in committees for the development of the MEAP tests managed by the state government. Her approach to her job, especially as it related to assessments, seemed influenced by her business background. She repeatedly stressed the importance of being accountable for teaching what the state standards require.

Classroom

Christy's classroom is neat, with no student work on the walls. She has a large collection of stuffed animals of various types that cover a bulletin board and cabinets. A sign on her whiteboard says, "KIDS AT WORK." Along one side of the room are microscopes and textbooks on carts to be used when needed.

[18] Moodle is an open source technology that at the time of this study had a developing set of worldwide partners. More information is available at www.moodle.org.

[19] Blackboard Corporation is a publicly traded company with a range of software solutions for educational organizations. It began in the higher education market and at the time of this study was expanding into the K-12 market. More information is available at www.blackboard.com.

Photo 5.9 - Christy Connolly's classroom front board


Photo 5.10 - Christy Connolly's side whiteboard with stuffed animals and beanie babies

Teaching and Assessments Philosophy

Christy seemed committed to making the experience of learning science special for her kids so it might serve as a gateway to their futures. Her decision to become a teacher was similar to the description of a vocational calling, although she didn't use those words. Her approach to teaching seventh graders was influenced by her understanding of their stage of life, and she seemed especially concerned with making science a good experience for her female students, mentioning this on several occasions.

Christy separated the students by gender on different sides of the room, saying she wanted to create an environment where students could “enjoy science.” She talked about making her classroom a safe place for kids to like exploring the natural world without being considered “weird” or feeling self-conscious for wanting to, for example, spend a lot of time examining the legs of insects. Christy mentioned teaching science at this age as an *organic process*. She stressed that science was something anyone could do, saying, “We are all biological creatures.” While she offered some basic principles that guide her instruction and assessment, she also described her approach as involving a good deal of improvisation and experimentation. Many times in response to questions about what she might do in certain circumstances, she described an evolving process where she might change her mind based upon quiz results or, as she said, “depending on what I want to do with them [students].”

Christy is also committed to students doing well on formal tests. While she does give students tasks that she believes are achievable, especially early in the year, she holds them responsible for their performance on assessments. She states in Transcript 5.6 her belief that students need to be able to succeed in high stakes testing situations.

Transcript 5.5 - Christy Connolly discusses students’ options for succeeding on assessments.

1)	PP: How important is it for a student to have multiple ways to succeed. I mean is that something that you find that ... OK if they are not going well on tests that they are going to be able to get it from their journal or or their labwork or does all go together?
2)	CC: I think it does. I mean I really think to me if they at some the test should be something that is passible to anybody.
3)	PP: Umhum… it shouldn't be out of reach?
4)	CC: If they're doing their their work, doing the class work and asking good questions, focusing and paying attention and can maintain an interest in studying, then I think a test is a great way for people understand of people showing what they know. I mean I think what we have now, we had it last year, but I don't know if we talked about it. Y'know we have these process grades.
5)	PP: uhum we did.. process and product.
6)	CC: yeah process and for math it is even more stringent than it is for science. Because for math you have 70% of their grade is based on assessment .. uh which is y'know WOW for seventh grade. It really puts a lot of pressure on them especially because everybody has to get a C y'know you have to pass with 70% ,um, it makes it very very difficult. And science to me for science is very organic, and I think the kids can get that.
7)	PP:Yes.
8)	CC: Y'know it's um if they are paying attention even if they are not really interested and they are doing the work, y'know looking at the website and going through the

Transcript 5.5 - Christy Connolly discusses students' options for succeeding on assessments.

process maybe making a little PowerPoint slide or um I have them like I like them to make an index card card of something that we put the index cards together into a poster y'know just that kind of stuff that I can look at and go *Yeah you got it that's very cool.*

9) PP: Like this kid who does the the Venn diagrams? [see document 5.5.c]

10) CC: Yeah I count those as assessments.
So I guess I think they should be able to show in multiple ways that they know stuff.
Um but let's be fair—let's be honest The major way people get places .. like **college** is by showing things on a test.
And there is no getting around that y'know I think I would be uh remiss in my **duties** if I didn't explain to students is that they need to know how to take these tests.
And that they learn that this is the way you're going to be looked at.
Y'know you can can once you get in you can do whatever you need to do.
But you won't be able to do it if you can't get in or get in or get out as the case may be

11) PP: Even those of us well past middle school need to get out.

12) CC: [Laughter] Right exactly, so y'know that's a big to me that's a big part of seventh grade. I don't I think they really need to be shown and told that this is important and you need to be good at it and if you are not good at it we've got to figure out a way to get you good at it.

Christy does not rely much on the textbook, considering it a resource that is available for certain purposes. Both she and Ben Raminskis were provided with the same Prentice Hall textbook series that Faith Churchill has. However, she and Ben took a different approach with the published books. She said she has often assembled course materials from online resources and delivered them to the kids via course Moodle web sites, in essence creating her own chapters. When asked about this, she said she was not writing a book, but she stressed that she used the term "chapter" to denote a progression, where the more common term "unit" implied that the instructional sequence could be easily altered.

<u>Assessment System</u>

Christy discussed her approach to assessment improvisation based on some general principles. Using Moodle for much of the class work, she begins each unit with a pre-test (called a quiz on Moodle) and then ends with a post-test also on Moodle. In between, she might give several quizzes, depending on how the students progress. She said that she gradually increased the demands and expectations on the kids as the year progressed, and she also frequently used informal tasks. Each chapter includes a major project or paper where students can use a variety of forms of expression to make their ideas explicit. She also has the students maintain a folder/journal of their work. She offered to make available her course website, but several attempts to view it failed, and my request for the guest password were never returned.

Table 5.5 - Assessment activities for Christy Connolly

Assessment Activity	Activity Frequency	Duration	Activity Relationships	Activity Source	Comments
Pre-test called a quiz	Beginning of unit	30 minutes	Same content as post test	Teacher assembled from variety of sources	Quiz is term used in Moodle for any assessment
Unit test (a post-test quiz)	End of unit	1 period		Teacher	Moodle
Quiz inter-mediate	7-10 days	15-30 minutes	May include questions for post-test	Teacher	Moodle
Lab report	Monthly or more frequent	2-3 periods	Vary in complexity depending on unit and position in year.	Teacher	
Projects	End of most units	1-2 periods	Summarize/synthesize	Students create to rubric	Various formats rubric graded
Journals	Weekly	1 period total	Project related	Students create	

* examples provided.

When considering the representation of her assessment approach in Figure 5.5, it is important to remember Christy's organic approach and her improvisation. As with Betsy Dearing, Christy's assessment approach could vary throughout the year, and she made it clear that she might alter that approach for each group of students.

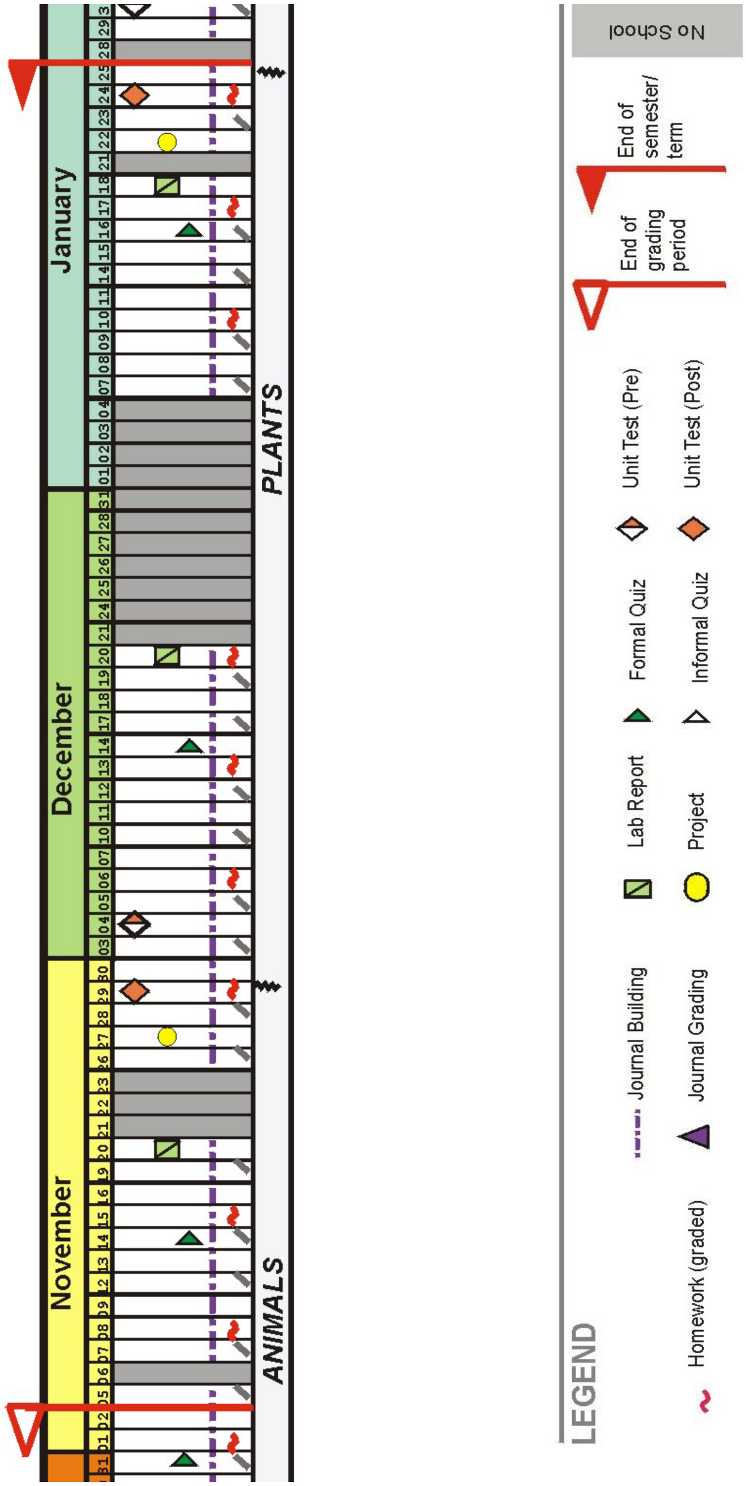
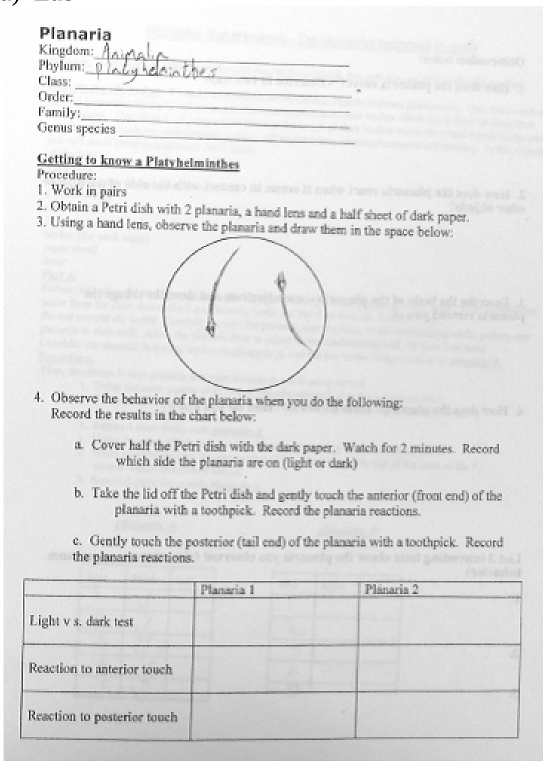


Figure 5.5 - Illustration of Christy Connolly's assessment system over time

a) Lab



Planaria
Kingdom: Animalia
Phylum: Platyhelminthes
Class:
Order:
Family:
Genus species

Getting to know a Platyhelminthes

Procedure:

1. Work in pairs
2. Obtain a Petri dish with 2 planaria, a hand lens and a half sheet of dark paper.
3. Using a hand lens, observe the planaria and draw them in the space below:
4. Observe the behavior of the planaria when you do the following: Record the results in the chart below:

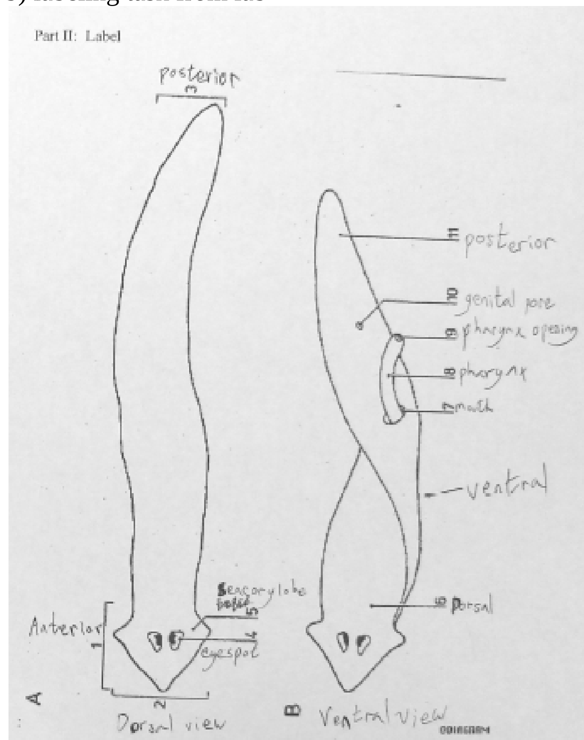
   a. Cover half the Petri dish with the dark paper. Watch for 2 minutes. Record which side the planaria are on (light or dark)

   b. Take the lid off the Petri dish and gently touch the anterior (front end) of the planaria with a toothpick. Record the planaria reactions.

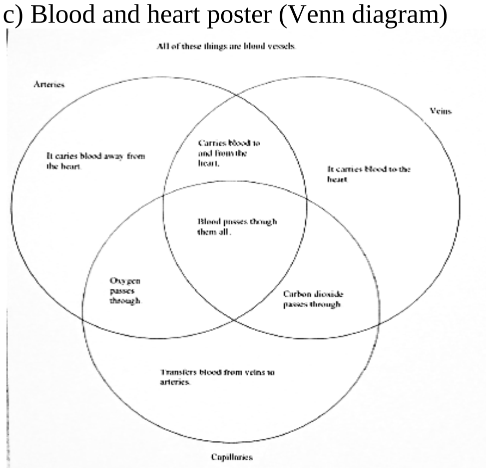
   c. Gently touch the posterior (tail end) of the planaria with a toothpick. Record the planaria reactions.

	Planaria 1	Planaria 2
Light v s. dark test		
Reaction to anterior touch		
Reaction to posterior touch		

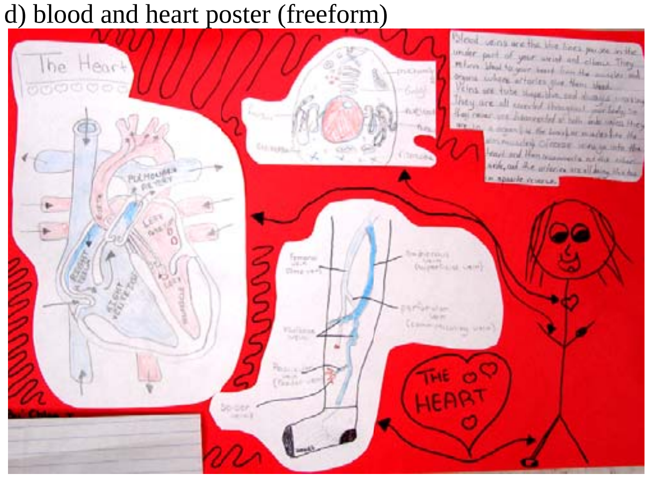
b) labeling task from lab



c) Blood and heart poster (Venn diagram)



d) blood and heart poster (freeform)



Document set 5.5 - Assessment texts from Cathy Connolly's classroom

Work with the State Test Development Process

Christy has participated in a number of committees for the state working on MEAP development (the formal process is discussed in more detail in Chapter 8). She indicated that the experience and being part of the process were important and beneficial for her. She originally became involved, she said, when she heard other teachers complain about the MEAP. She decided to find out more about how it was developed and states in Transcript 5.6 why she would recommend it for other teachers.

Transcript 5.6 - Christy Connolly discusses the experience of working with the state test development process.	
1)	PP: Do you think the experience, whether or not it would be valuable for the state, for new teachers, for young, for different types of teachers would it be a valuable experience to go and be part of those committees?
2)	CC: Oh YEAH. Absolutely.
3)	PP: Why?
4)	CC: Um it well first of all you become very familiar with the objectives and the standards.
5)	PP: OK.
6)	CC: You know exactly what it is that the state wanting you to be teaching. And if maybe a committee is isn't the right place to do it .. but at least have some sort of way to know for teachers who are new into the profession have the ability or have the .. the resource that they can meet and sit down and um look at that. And I don't know maybe it's done in districts in other districts.. but it would be a great idea for the state especially if the state is trying to re-form the science with the new standards are coming out.

The Case of Ben Raminskis, Crimson Middle School

Ben Raminskis had been a science teacher for 28 years and taught both seventh and eighth grade science in the room next to Christy at the time of this study. He had been active in the MSTA in the past, but had let that membership lapse. He was, however, an active member of the National Science Teachers Association (NSTA). He is the official department head for the school and a member of the school's management committee. He also managed the school science fair.

Ben is a bearded, latemiddle-aged man with a quiet and patient demeanor. He was a willing participant in the study. He joined the study serendipitously. I first met him at 7:00 a.m. in Crimson Middle School. I had come to meet with Christy and found on that day that there was a special event in one of the classrooms where students were operating a "café" for the adults. While these teenagers were taking orders for juice and bagels, Ben sat down at my table, and upon hearing of the study agreed to participate. He expressed curiosity and interest and volunteered his experiences with developing the state standards and tests some years earlier. I came to learn later that, perhaps unsurprisingly for a teacher in a school so near my research university, he knew about many of the faculty members at the University of Michigan. Many doctoral students had likely passed through his school before.

Classroom

Like most classrooms in Crimson, Ben's was neat. There were no projects on the walls or other examples of student work visible. Rather, he displayed various artworks, including Asian prints and Ansel Adams photographs on the walls. Like Christy's room, Ben's had two lab stations with sinks and gas outlets in the middle. Along the back were more lab stations, and along the sides were cabinets. He did not separate his students by gender.

Photo 5.11 - The front of classroom with University logo on whiteboard.

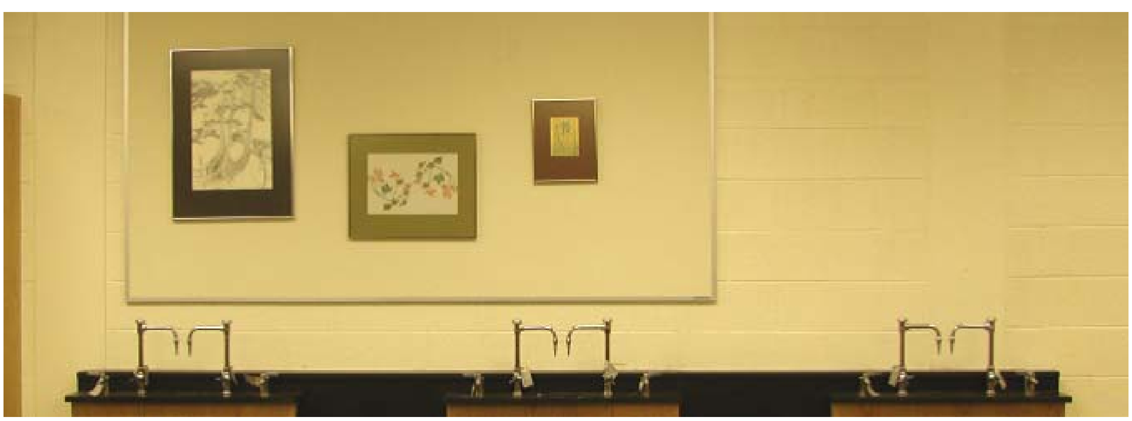


Photo 5.12 - Back of Ben Raminskis' classroom

Teaching and Assessments Philosophy

One of the signature features of Ben's approach that came across in our discussions was the concept of gradually withdrawing supports and constraints from students as the year progresses. He termed it “fading.” Early in the year he would provide the students with a lot of structure and then, as the year progressed, he would increase his expectations that they would do more for themselves. He did not describe a dramatic shift or change in expectations, but rather a consistent progression. This approach was similar to what his colleague Christy described as being like a volume knob on a stereo that only went up as the year went on. He also discussed how this shift in support and expectations could continue into the next year if, as he said, he is “lucky enough to have them for eighth grade also.”

Ben did not rely heavily on the textbook. Like Christy, he reported using it as an occasional resource. His instructional emphasis seemed to be balanced on scientific method, the development of scientific representational skills, and traditional science classroom expectations for teamwork and lab skills. While he indicated he did little to foster scientific literacy, his assessment tasks included having students draw and in some cases invent graphics, as well as use scientific terms. He also had them write lab reports to show competency with collecting and presenting data.

He described the current period of standardization and the job of teaching science in balanced terms and observed that, while he had been motivated to study science by his own experience with real phenomena and engagement with the outdoor world, these options were less available to students today. Authentic science, Ben believes, is being supplanted by more focus on shallow and artificial tasks and less focus on direct experience. As an experienced teacher he seemed to favor a conversational approach, describing one method of monitoring student progress as just talking to them to find out “what’s cooking.”

Assessment System

Ben was efficient in his use of words. And, like his colleague Christy, he showed only a few examples of assessment instruments. However, he was able to describe his system quickly by instrument and type of task. He described his approach as “traditional,” including tests and quizzes, and he described it as being almost totally created by himself. He used a variety of task types and instruments in a regular and repeated pattern, with two quizzes that are followed by a test. The quizzes and tests shared similar features, including various tasks. While he described his approach as traditional, the instruments he used exercised a range of semiotic tasks.

Table 5.6 - Assessment activities for Ben Raminskis

Assessment Activity	Activity Frequency	Duration	Activity Relationships	Activity Source	Comments
Quiz *	2 times per marking period	10-20 min.	Similar in form to test. Unit content.	Teacher	
Test*	Once per unit	Full class	Similar in form to test. Unit content.	Teacher	Sometimes prepares a prep
Project/ report	Once per marking period	2-3 classes	Could be used as the test.	Teacher	
Lab report*	Twice per marking period	2-3 classes	Unit content	Kids develop from teacher model	
Pop quiz	Occasionally, early in the year	10 min.	Unit content	Teacher	

* - examples provided.

Figure 5.6 illustrates how part of Ben's year might progress. Ben taught four units per year. Each one generally had two quizzes and ended with a test. He sometimes gave his students a pop quiz and said he sometimes had a test prep session if he felt it was needed. As with Christy, he expressed that he improvises and makes decisions about what is needed and adjusts accordingly. He wrote on the draft document that he reviewed for this study: "Phil, this is pretty accurate! I really sort of schedule these things informally." An important part of his year involved the science fair that he managed for the school. Even though Christy's students would presumably participate in the science fair, she never mentioned it in any of our discussions. Ben's approach contrasted with Paul Bond's. It was not only less frequent in terms of textual tasks and its feature of a core project area, but it was also less fixed in terms of what occurred on any given day.

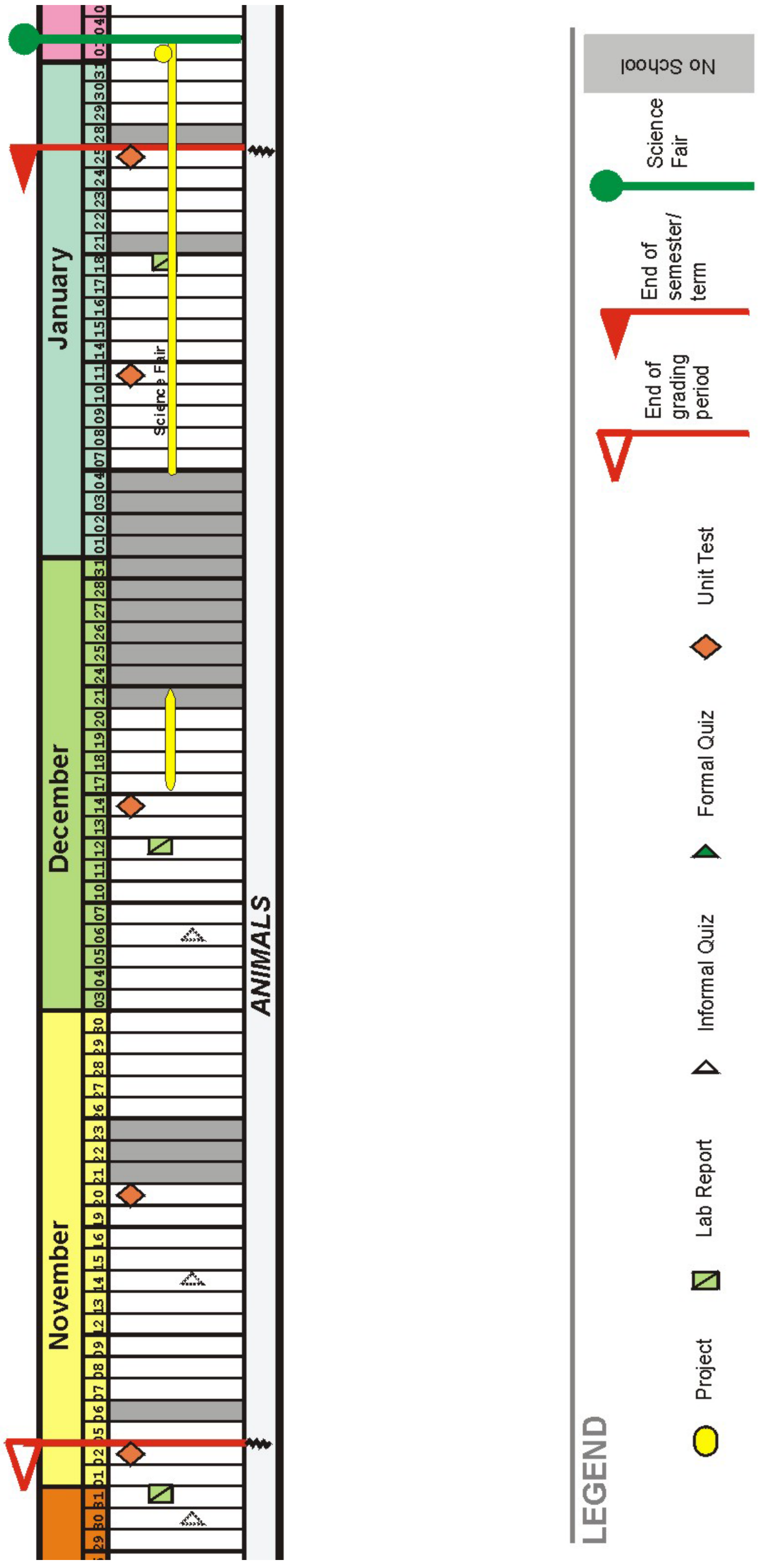


Figure 5.6 - Illustration of Ben Raminskis' assessment system over time

a) Test

VERTEBRATES TEST

MULTIPLE CHOICE. Place the letter of the correct answer on the space provided. (1 pt each)

1)__D__ All vertebrates have a) exoskeletons b) scales c) hair d) bony skeletons.

2)__C__ Which is best suited for life in water? a) toad b) frog c) goldfish d) salamander.

3)__D__ Which is not a vertebrate? a) snail b) squid c) fly d) non of these are vertebrates.

4)__A__ Which group of vertebrates are warm blooded? a) mammals b) amphibians c) reptiles d) fish

5)__A__ A shark would be classified as a) a cartilagenous fish b) a jawless fish c) an amphibian d) a reptile

6)__C__ What characteristic do fish and amphibians have in common? a) scales on skin b) lungs for breathing c) must lay eggs in water d) warm blooded.

7)__C__ Immature (young) frogs breathe through their a) lungs b) gills c) skin d) mouth

8)__B__ Which is not an amphibian? a) frog b) vampire bat c) newt d) non are amphibians.

9)__C__ Bony fish can float at different levels in water because they have a) backbones b) fins c) swim bladders d) gills

10)__D__ Which creatures are not reptiles a) dinosaurs b) snakes c) lizards d) toads

11)__B__ Reptiles all a) have gills b) lay their eggs on land c) have hair d) have smooth, moist skin.

b) Test with essay portion

5)__True__ When a bird flies south in the winter it is migrating

6)__False__ When comparing an alligator and a crocodile, if you can see the teeth with the mouth shut you are looking at an alligator.

7)__False__ All reptiles give birth to live young.

8)__False__ Snakes use their tongue to sting.

9)__False__ Newts and Salamanders are really fish with tails.

10)__False__ Snakes, Lizards and Turtles are all amphibians.

SHORT ANSWER, ESSAY

1) Tell two similarities (compare) and two differences (contrast) between amphibians and reptiles. (5 pts)

2) Name (and explain) three key characteristics of birds. (5 pts)

c)Quiz

ARTHROPODS QUIZ

MULTIPLE CHOICE Place the letter of the correct answer on the space provided.

1)__B__ Characteristics of ALL arthropods include a) 8 legs b) an exosleleton c) wings d) gills for breathing.

2)__C__ Bilateral symmetry means that all arthropods a) have at least 6 legs b) have several body parts c) have halves that are identical if you split it down the middle d) have at least two laterals.

3)__A, C__ A crustacean would be a) a lobster b) an ant c) a hermit crab d) a millipede.

4)__D__ Arthropods called "trilobites" are unique because a) they live in the Huron river b) they all had at least 40 legs c) they were able to make their own food (photosynthetic) d) they are extinct.

5)__B__ The three body parts of an insect include a) head, tail and legs b) head, thorax and abdomen c) head, belly and but d) abdomen, appendix, and intestines.

6)__A__ Myriapods are a group that includes a) centipedes and millipedes b) ants and wasps c) fiddler crabs and hermit crabs d) lobsters and crayfish.

7)__D__ The really big claw on a Fiddler Crab is a) found on males only b) called a cheliped c) used to attract females d) a,b,c are all correct.

8)__A__ Which characteristic applies to spiders only? a) venom b) compound eyes c) paired legs d) 3 body parts.

d) Diagram and labeling task

DIAGRAM Make a pencil sketch of an Arthropod. Use a key and label at least 10 parts.

LABELS

1)

2)

3)

4)

5)

6)

7)

8)

9)

10)

Document set 5.6 - Assessment texts from Ben Raminskis' classroom

The Case of Valerie Jones, the Kit Teacher at Hardy Middle School

Valerie Jones (VJ) was teaching seventh grade science in Hardy Middle School and, at the time of this study, she had been a science teacher for 10 years and active in MSTA for eight. This case study is unique in that in some ways Valerie Jones represented several teachers. Hardy Middle School uses a science kit program, in which materials are shared among a group of teachers from a central resource pool. Hardy has four seventh grade science teachers plus a special education teacher. All five seventh grade teachers participated in this kit program that is administered by their small district's science center, together with teachers at some local private schools. Valerie indicated that of the schools that used the kits, Hardy was one of the best equipped. It did not need the kits in order to operate a science program, whereas some of the other schools really needed the supplies and materials the kits provided. The result of this program is that only two science teachers in the school can teach the same topic at the same time. Some kits contain commercial textbook materials. Some use a special supplement called a reader.[20]

Photo 5.13 - Science kit materials ready to be shipped to the next school

Valerie was also the only teacher who was interviewed just once. On the one day I met with Valerie, she was in a shared office area rather than in a classroom, because on that day a student teacher was working with her class. Two other teachers, Duane Sprocket, who teaches eighth grade science, and Angela Dubois, who previously taught seventh grade science and recently started teaching eighth graders, were also in this shared office. They participated in much of the interview with Valerie. The result was

[20] This term "reader" is also used in the *Investigating and Questioning Our World with Science and Technology* (IQWST) program for the same type of textual support.

an interview that yielded more of a description of the school's program and how many teachers worked with it than a description of Valerie's specific practice. The member checking, however, focused on Valerie, and the representations of the assessment system in Figure 5.7 below relate to her particular implementation of the kit program rather than to a general description of how all teachers in the school use the kits.

The school differed somewhat from others in the study in another way because of its personnel system and a reform effort that was underway when I visited. The district had a strong union and teachers frequently changed positions in an annual process that they called "the bid." Within the bid system, any teacher could apply for any open position, with seniority being given a high priority in the selection process. Under this system, a teacher could lose a position not for lack of qualifications, but because another teacher won it in the bid. This bid system was discussed by all of the Hardy School participants, including some who will be introduced in the next chapter. The school itself was reported to have a history of adversarial relationships between administration and faculty. The school year of this study was the principal's first year in that role and in the school.

Because of the kit program and political climate, perhaps this case should be viewed primarily for its contribution of a kit teacher perspective. Within this system, teachers have less control over large blocks of instructional time and short gaps in materials that they manage on their own.

<u>Classroom</u>

Valerie's classroom had student desks arranged in the middle of the room behind four large lab stations. One wall of the room was lined with supply cabinets. In one corner of her classroom was a timeout section called "the spider's web," and the bulletin board was covered with sticky notes being used to depict a chemical reaction.

Photo 5.14 - Classroom and spider web corner from Valerie Jones' room

Photo 5.15 - Bulletin board depicting photosynthesis reaction in Valerie Jones' room

<u>Teaching and Assessments Philosophy</u>

By providing supplies and texts to support the lessons, the science kits provide a structure and method overlay for the teachers. The teachers then have less freedom to create their own classes. In elementary schools, kit programs are common; one reason is that they can compensate for teachers who lack science knowledge. In addition, the bid system used in the district may have helped fostered a high turnover and lower ownership of the classroom. Valerie, however, is a senior teacher who does have science knowledge, so she did not fit the same mold of many elementary teachers who need to provide some science instruction without the background.

Valerie and her co-informants described themselves as “big project people” and showed many examples of student posters and constructions. Homework was not stressed in her curriculum, and Valerie indicated it was used for cases when the work was not completed in the class.   In Transcript 5.8, Valerie and Angela discuss their approach to projects.

Transcript 5.7 - Valerie Jones (VJ) Angela Dubois (AD) discuss projects

1) VJ: We're big project people so for every unit we have projects
2) AD: Every kit we had a project so that if they didn't do well on their test there was another big ticket as far as points are concerned item that they could do well on.
3) VJ: And we developed those together so that there were rubrics and guidelines and that kind of thing so we were real consistent with our kids ended up with the same kind of product.
4) PP: But that's not going to be actually in the kit but that is what you decided as a department to do?
5) VJ: Well, I think we gave them to [district science center manager] Debbie to put in the kits but they are optional.
6) AD: The teachers can do what they want as well. ...
   It is probably easiest with the electricity unit to go through. We did they could create a flashlight, they could create a house, they have chose to make a conductor/insulator they could have chose to write a children's book.
7) PP: To write a children's book?
8) VJ: Uhhmm and for each different whatever they chose there was different criteria and different directions.
   But the same rubric was used no matter what it was that they chose to do.
   So there was like a project list I want to say there was five different things they could choose to do.
   And they could choose to do it based upon what they wanted to do.
   The writer, the children's book would be for those who would like to write and that.
   Most of the other ones were hands-on things.
9) PP: Does this help like with the the boys and the girls? Are the girls doing more writing and boys doing more flashlight building?
10) VJ: Not necessarily, but that was our thought process well those that can't those are more hands on oriented rather than .. a writer or someone who can sit and can think about that ... or someone that wants illustrate something or whatever. That was our thought process but no not necessarily.
   The people that that will tackle the writing are mostly girls that will do that.
   But most most of the kids at Handy .. at this school will do the projects hands on project hands down.   I would say .. probably 98% of them would do the hands on pick one of the hands-on projects, and a very small maybe two or three kids every year would do the writing.

Assessment System

The assessment system used in the kit program includes a pre-test and a post-test as well as some quizzes, labs, and activities in the middle of the unit. Each teacher can choose to use only some aspects of the kits, including the assessments, and/or to add their own material/assessments into the instructional sequence. Valerie, like Christy, said that she does both the pre-test and the post-test and gives the kids quizzes every two to three weeks, although only a few samples of assessment instruments were provided. The science coordinator for the district, Debbie Huston, confirmed that the kits contain these assessments. However, I was unable to locate pre/post-tests and I was not able to see any reports indicating that the results of these assessments were systematically compared.

Table 5.7 - Assessment activities for Valerie Jones

Assessment Activity	Activity Frequency	Duration	Activity Relationships	Activity Source	Comments
A. Bell-ringer	Daily	Minutes	Could be test questions	Valerie	
B. Lab report	Weekly	50 Minutes		Kits/Labs/ books	Build understanding
C. Quiz	Approx. 10 times per year	50 Minutes		Kit + Valerie's additions	
D. Unit pre-test	4 times a year	50 Minutes	Unit content	Kit + Valerie's additions	
E. Unit post-test*	4 times a year	50 Minutes	Unit content	Valerie	
F. Project/poster*	4 times a year, in gaps of kits	1-2 weeks. A little each day.	Often includes material not in kit		

* - examples provided.

When considering the illustration of Valerie's teaching practice shown in Figure 5.7, there are two important features to note. First, Valerie provided this information based upon reviewing her own calendar for the previous year, so her placement is like a documentation of past practice. Second, since Valerie uses a kit program, the timing of unit boundaries could be regulated by the district kit schedule, but this seems to not always be the case. Even though the Matter Matters unit material was scheduled to return to the rotation in mid-January, Valerie continued teaching that unit (probably without the kits) and did not conduct the post-test until early February.

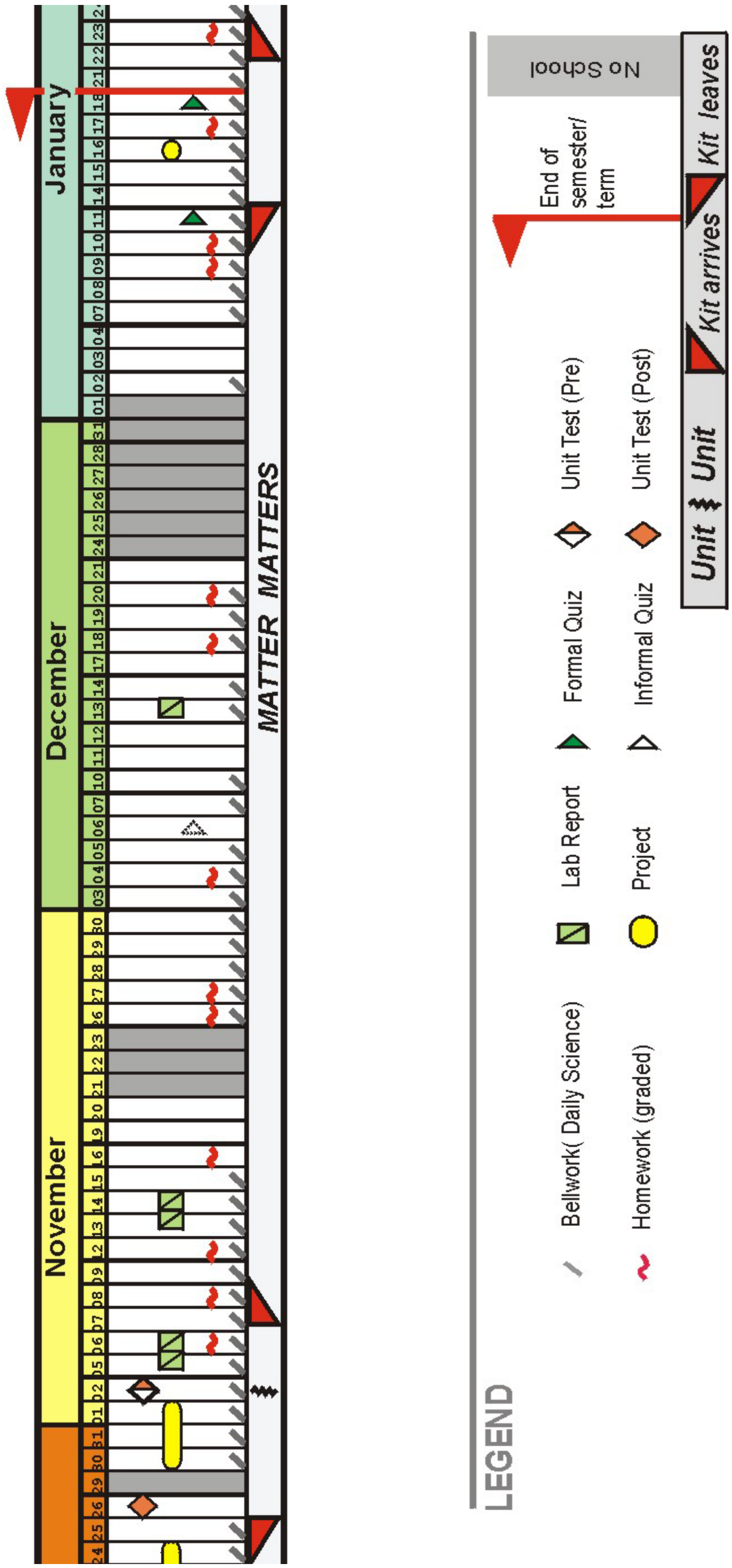


Figure 5.7 - Illustration of Valerie Jones' assessment system over time

a) Test

Core:

Matter Test

Circle the best answer

1. Steven is riding the bus to school on a cold January morning. He breathes on the window of the bus and notices the window starts to fog up. This is an example of
   a. evaporation
   b. condensation
   c. expansion
   d. contraction
2. Mr. Stevens set a pan of water on the radiator in his house. Three days later he noticed the water level was greatly decreased. Therefore, what process did Mr. Stevens conclude took place?
   a. condensation
   b. expansion
   c. evaporation
   d. condensation
3. Samantha put an ice cube tray filled with water into the freezer. The following day she noticed she now had a full tray of ice cubes. Several months later she needed some ice cubes for her soda. When she opened the freezer, she noticed the ice cubes were much smaller. What caused the ice cubes to shrink?
   a. sublimation
   b. condensation
   c. contraction
   d. thermal expansion
4. Marcus lit a piece of newspaper on fire and set it on a tray as part of a Science experiment. He observed that all that was left was a pile of ash. What can you conclude from this?
   a. A physical change occurred
   b. A phase change occurred
   c. No change occurred
   d. A chemical change occurred
5. Think about the ash that was left over from the burnt piece of paper. What would you call this part of the experiment?
   a. reactant
   b. product
   c. variable
   d. mineral deposit
6. Which statement best describes the motion and arrangement of molecules in an ice cube?
   a. Vibrating in place with a regular repeating arrangement
   b. Random motion and packed closely together
   c. Random motion and widely spread apart
   d. Vibration in place and slightly separated

b) Project assignment sheet

Name:
Core:

This sheet must be turned in with project

Electricity Project

House Grading Sheet

Design and Construction
/5 Project design mimics a real home.
/5 At least 3 rooms.
/5 House is well constructed, sides are together, dividers for walls are sturdy.
/5 Wiring and circuits mounted securely.

Electrical
/5 House is wired in each room.
/5 There is a switch to turn power on and off.
/5 Circuits work properly.
/5 Power is supplied by a battery.
NO PREWIRED KITS

Neatness
/5

Rooms
/5 Lights mounted in wall, fireplace, ceiling or lamp.
Wires aren't visible – hidden in walls.
/10 2 way switch(es)

Total Points Possible
60/55
$10 limit on materials
excluding batteries

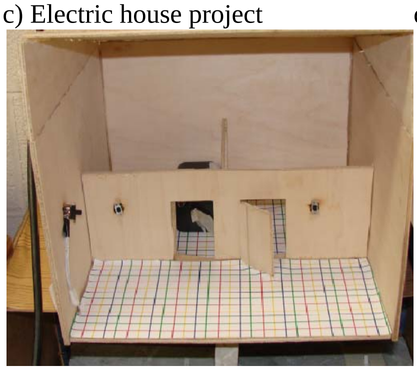
Liquid Conductor Grading Sheet

Circuit
/10 Circuit works properly
/10 There is a switch to turn the power on and off.

Liquids
/10 7 liquids were tested with the circuit.
/10 The project contained conductors and insulators.
/5 Liquids in sealed container.
/5 All containers were labeled with the name of the liquid.
/10 Project was neat and cleanup was done after presentation.

Total Points Possible
60/55
$5 limit on materials

c) Electric house project

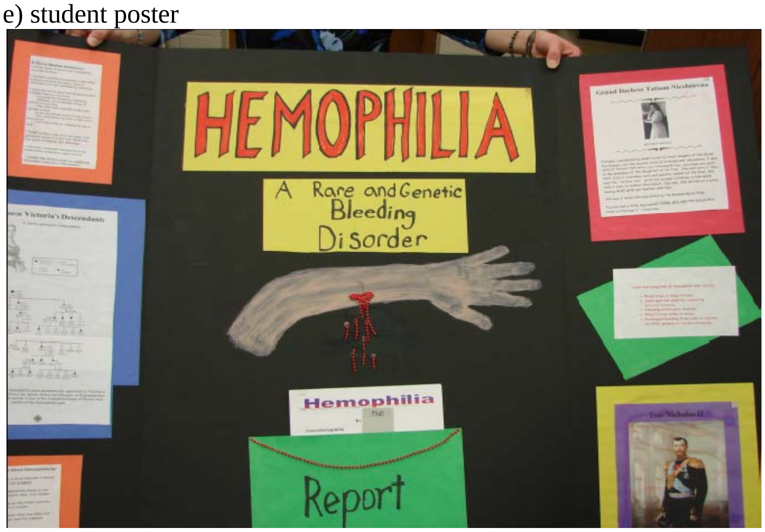


d) House project grading rubric

House Grading Sheet

/15		**Design and Construction:**
	/5	Project design mimics a real home.
	/5	At least 3 rooms are present.
	/5	House is well constructed; sides are together, dividers for walls are sturdy.
/20		**Electrical:**
	/5	House is wired with electricity in each room.
	/5	There is a switch to turn power on and off.
	/5	Circuits work properly.
	/5	**The power supply is by** battery.
/5		**Neatness:** Project is neatly done.
/10		**Rooms**
	/5	Rooms have (pretend) flooring (carpet, tile, wood, etc).
	/5	Walls have color
/0		**Extra credit:** Furniture, people, switch in each room, additional rooms, other.

e) student poster



f) Beadwork on poster

Document set 5.7 - Assessment texts from Valerie Jones' classroom

Cross-case Analyses of the Case Teachers

These seven portraits of practice illustrate some of the ways these teachers deployed assessment instruments as coordinated, deliberate systems that occurred in often regular intervals. These systems are individual, however. And while all seven teachers teach the same subject in the same grade, the specific nature of their assessment approaches often varied. In this section I look across these seven cases and attempt to identify patterns. A conceptual perspective I use across these analyses is that of classroom practice as a conversation between a teacher and students (Yinger, 1990). Yinger's dialogic framework describes instruction as occurring within a cyclical process between teachers and their community of students. Each of these cross-case analyses in some way provides additional details about that conversation.

Analysis of Task Structure

This short analysis looks at the task structures students may encounter in their classroom assessments. In this analysis I will explore the type of work students are given to do in these classrooms, asking whether, across the seven individual systems, there were any common patterns to the types of tasks students are asked to perform. Table 5.9 shows the tasks teachers reported giving their students. It was first developed from analysis of the teachers' documents and then confirmed and refined in the member-checking process. This classification of tasks is external and shows that, with a few exceptions, the type of work students do in these seven teachers' assessments fall within similar classifications.

This finding illustrates that for these teachers, the conversation of practice often involves textual tasks which reinforce a multimodal literacy that can be part of contemporary science classrooms (Lemke, 1998). In addition to exercising the range of two-dimensional semiotic systems in their assessments, some teachers encouraged students to express themselves with more modalities. Some used tactile elements, as illustrated by Valerie Jones' projects and the posters in Christy Connolly's and Jim Heinrich's classes. Some, including Betsy Dearing, Paul Bond, and Faith Churchill, also used the building as a textual space. Paul used the ceiling[21] and Betsy used classroom walls, in both cases to illustrate issues of phenomenological scale. Similarly, Faith used her cafeteria's wall as a gauge to measure the height of hot air balloons. Christy Connolly added into this study electronic assessments on Moodle that in the future will likely include animations and simulations. It is, then, possible to see the conversation between teachers and students that this study describes occurring across a broad spectrum of modalities.

Table 5.8 - Summary of task types for the seven case study teachers

Task types in assessment systems		**PB**	**FC**	**BD**	**JH**	**CC**	**BR**	**VJ**
Choice	True false	X	X		X	X	X	X
	Multiple choice	X	X	X	X	X	X	X
Verbal	Fill in blank/matching	X	X	X	X	X	X	X
	Definitions		X	X				
	Short answer	X	X	X	X	X	X	X
	Extended Response	X	X		X	X	X	X
	Essay		X			X	X	X
Diagrams	Labeling	X	X	X	X	X	X	X
	Reading	X	X		X	X	X	X
	Building	X	X		X	X	X	X
	Inventing	X	X			X	X	X
Evidence/Data	Interpreting	X	X	X		X	X	X
	Creating	X	X	X		X	X	X
	Explanations	X	X	X		X	X	

(Column group heading: - - - Teacher - - -)

Analysis of Function: Mediational Purposes and Boundary Crossing

While the previous analysis focused on form, this analysis focuses on function. It explores the purposes of these teachers' assessment practices by considering three perspectives: assessments as learning tools, assessment systems as indexes of implicit cognitive theories, and assessment information crossing boundaries.

The first perspective considers how assessments can be learning tools, as well as diagnostic tools. A few of the cases suggested that classroom assessments can support learning in two types of recipients: students and teachers. For students, the activity of

[21] There is evidence that this ceiling display was probably done by the 8th grade students and not seventh graders. Still, it is part of Bond's classroom approach, and the Periodic Table of the Elements is from the seventh.

working on an assessment seems to aid learning. For example, Betsy Dearing gave a detailed account of how assessments could help students to solidify their knowledge in confusing areas. She said that in at least one instance, she inserted an extra quiz into her lesson plan so that students would not be confused by similar sets of life science terms. - This is consistent with Betsy's attention to scientific terminology, as she was one of only two teachers who reported giving students word definition tasks, as shown in Table 5.8. All seven teachers were asked to rate the importance of giving students multiple opportunities to succeed. Five said it was extremely important, with the other two rating it as highly important. The retaking of assessments suggests some possibility that there is a general pedagogical value for students when they perform assessment tasks, although the specific details of this possibility were not explored in the study. In terms of teacher learning, Jim Heinrich talked about how assessments inform his teaching. He described them as similar to a baseball scorecard that lets the pitchers know "how many hits they have" with the students. There was no evidence in the study, however, that what he learned on his scorecard influenced the game he played or changed its parameters. Other case study teachers did not express a similar opinion, but respondents to the survey discussed in the next chapter did rate highly the question of whether they use assessments to reflect on their own teaching.

The second perspective in this analysis of purpose is related to the issue of assessment tasks as tools for learning. It considers how assessments may index ways of thinking of student cognition. This area involves the size of cognitive demand an assessment task places on students. Some of the case study teachers favored small conceptual units and some favored larger units in their assessment approaches. Jim Heinrich, Paul Bond, and Faith Churchill all used frequent, regularly spaced assessments. Faith, in response to my questions about what influenced her assessment development, said, "I have learned more from watching and experiencing how students take small steps of success." Jim discussed needing to assess frequently before the information "leaked out" of his students' heads. Paul Bond didn't discuss this aspect, but his assessment system and his use of an experiment worksheet given every day or two would suggest continual reinforcement of learning objectives with incremental additions. On the other hand, some teachers, such as Christy Connolly, Betsy Dearing, and Ben Raminskis, tended to use less frequent assessments and focus on more in-depth activities. Betsy designed elaborate homework project activities to teach students about nutrition and their bodies. Ben focused on the science fair, and Christy talked about the way she improvises with her assessments to meet the needs of different groups of students. This contrast highlights a possible alignment between assessment approaches and teachers' underlying belief systems that may reinforce each other. For if teachers assign and evaluate small tasks, may they be more likely to believe in cognition as an incremental process because they see students work in these small units? And might teachers who give larger, more integrated tasks similarly see the potential for cognition and learning in larger steps?

The third perspective on the purposes of classroom assessments is boundary crossing. "Boundary crossing" is a term that refers to information passing between different organizational units for some purpose. This area is an important topic in terms of systemic function, as systemic function often requires the coordinated interaction of different parts of the system. Across these seven cases, I was unable to identify any instances of classroom assessment information being used by others inside these teachers' educational organizations. At Crimson Middle School, Christy Connolly did discuss the discrepancies between grades and MEAP for math, and Ben Raminskis talked conceptually about the possibility that grades and MEAP might not agree. However, other than the Crimson school principal who monitored students' science project work, there was no evidence that anyone within Crimson at a broader level was using grade information or looking at specific classroom assessments.

The only cases of boundary crossing I found were not inside the school system, but between teachers and students' homes. Three teachers -- Paul Bond, Faith Churchill, and Valerie Jones -- discussed sharing specific, assessment instrument-based student performance information with parents. Paul sent home a sheet for parents to sign (Document 5.2.c). Faith had students lead a conference with parents where they used their journals to discuss their progress and plan for improvement. Valerie emailed parents weekly, advising them of upcoming student assignments.

Beyond those three instances of school-home connections, there were only a handful of instances where the teachers, or others in their schools that I interviewed, even considered seventh grade science classroom assessment information as having broader organizational use. One way I probed this was through a series of questions in the general interview that asked participants what the uses were, or could be, for assessments inside and outside the classroom (RQ5-9). Almost all participants who responded to these questions did not describe any connections between what is measured inside the seventh grade science classroom and other educational purposes. Two exceptions were Jim Heinrich and Ben Raminskis, who both speculated about the role that science classroom assessments might play in special education diagnoses. Jim wondered if classroom teachers could help expose home issues for students labeled as learning disabled. Ben made a brief mention of a relationship between special education and his assessment practices.

Christy Connolly provided a different perspective, stating emphatically that the purpose of her classroom assessments was to measure state benchmarks and that she didn't see any conceptual difference between her assessments and the state's MEAP test. Even though she is a math teacher, Christy did not describe opportunities to understand math and/or science arising from what occurs inside classrooms teaching the other subject.

With these few exceptions, classroom assessment was presented by these participants as a closed system.

Analysis of Temporal Structure

One of the theoretical influences I draw upon in this study is Lemke's (2000) description of the human semiotic systems as using timescales that relate to their ecosocial position or level. In Figures 5.1 through 5.7, I represented each of the seven case study teachers' practices with conceptual illustrations that spanned of a few months. These illustrations show that each teacher can have a unique meter for the types of assessed textual conversations (e.g.: quizzes, projects, tests, etc.) they engage in with their students. In the following analysis, I extend the comparison by showing all seven conceptual illustrations for an entire school year in Figure 5.8 (figure spans two pages).

Looking at the practices of these teachers over an entire year shows the individual nature of their assessment approaches more clearly. In this illustration I have also included a topical sequence schedule below each teacher's calendar. This identifies the topics each individual teaches as a year progresses. Not only is there no alignment of instruction across schools in the study, but also in the two schools where two teachers were studied, there was no alignment in their adjacent classrooms. As for the science kit teacher, Valerie Jones, the nature of the shared resources made alignment of instructional topics impossible across all classrooms in the school. However, her practice was not totally defined by the kit schedule, as she extended units, added material to one unit and started another unit late when she believed it was important. Perhaps most telling, while some teachers mentioned standards, not one of these teachers presented their curriculum plan in terms of the state benchmarks.

It should be noted again that Figure 5.8, like those before it, is a conceptual illustration. The school year represented in this figure is actually the 2007-2008 school year, which was the current year for the teachers at the time of the member checking process. In this process, a draft was sent to them based on a reconstruction from my notes, and they all returned the draft with various additions and corrections. As noted earlier, in any given year a range of unplanned events -- from weather to school emergencies to issues with students -- can cause teachers to need to adjust their plans. Still, three aspects of the temporal structuring of these teachers' classroom assessment practices that transcend individual expression do appear.

First, all of the teachers' practices use timescales shorter than the timescale of the state's annual MEAP cycle. Second, the teachers in the study operate within *multiple chronic systems* that include their district holiday and term schedules, their own topical sequence schedule, the material availability schedule (for Valerie Jones), and their own instruction-assessment systemic rhythm. Third, the teachers' individual rhythms could be expressed in terms of *signature characteristics*. These characteristics include timescale intervals, uniformity, and flexibility. The intervals relate to the time spans between textual activities. These teachers exhibited a range from the short intervals used by Jim Heinrich and Paul Bond to the longer intervals used by Betsy Dearing and Ben Raminskis. Uniformity is related to the likelihood that the intervals will remain consistent. Some teachers use consistent intervals, as Paul Bond and Jim Heinrich do, while others vary the pacing. Flexibility refers to the likelihood that temporal patterns may change. Some in this study (those with short intervals), including Jim Heinrich and Paul Bond, presented their approach as more concrete and less likely to be changed over the year, while others (those with longer intervals), including Christy Connolly and Betsy Dearing, indicated that they might alter their approach throughout the year as needed.

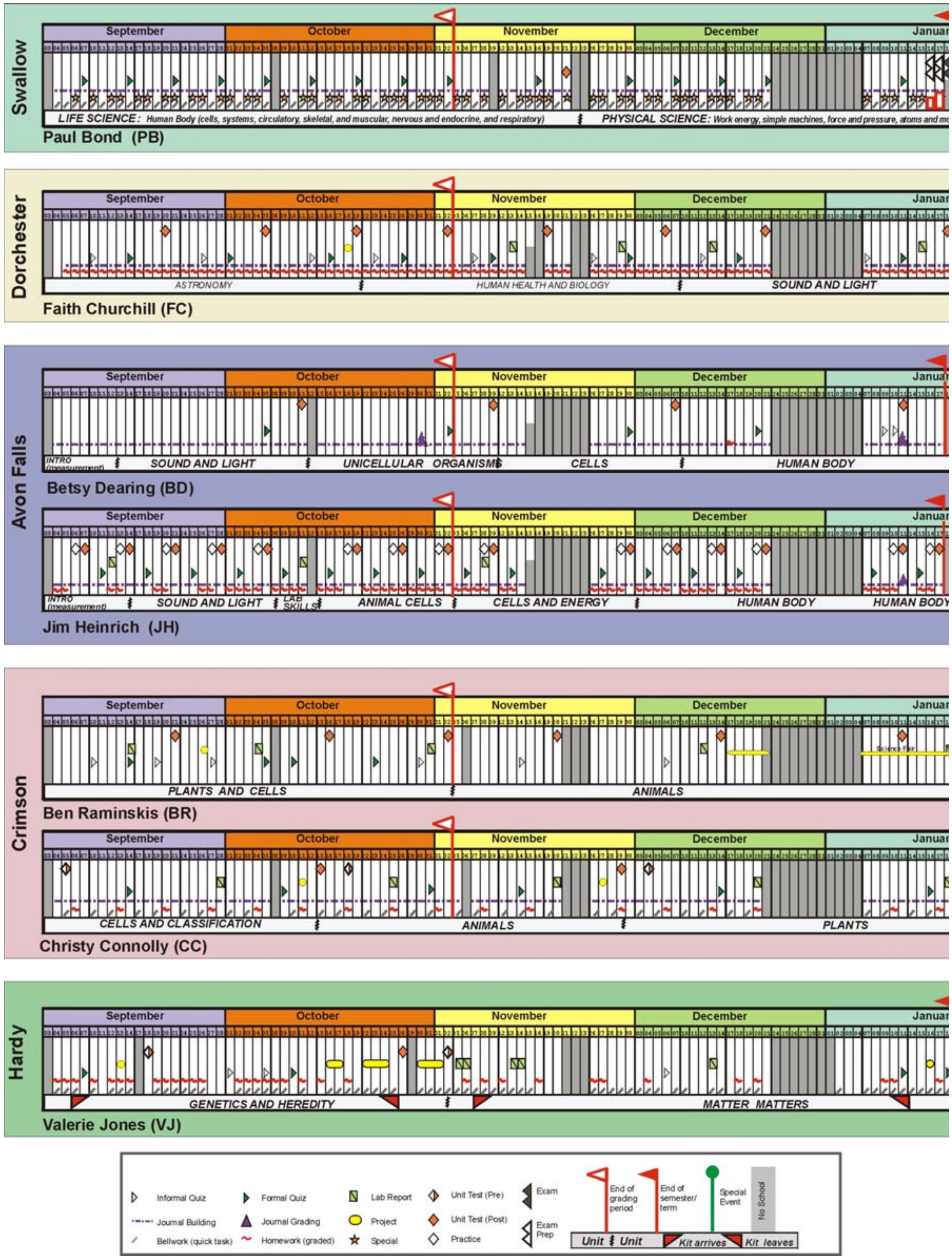


Figure 5.8 - Assessments calendar

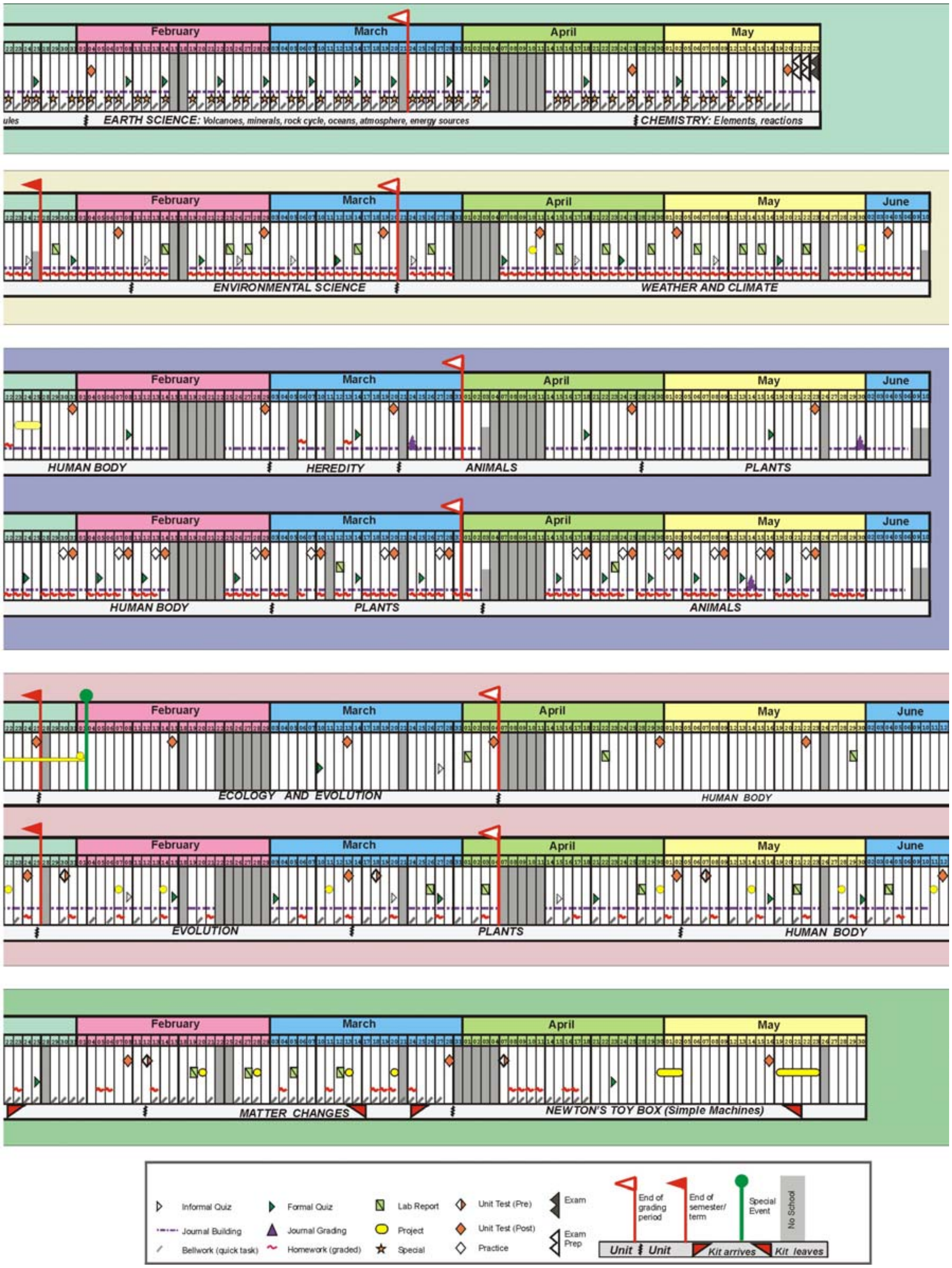


Figure 5.8 (part 2) - Assessments calendar

Analysis of Systemic Influence/Bricolage

These teachers largely created much, if not all, of their own assessment systems. Even those like Paul Bond and Faith Churchill who used textbook materials for tests, or Valerie Jones who was provided with tests, still created many other instruments. While these seven teachers differed in their approach to assessments, *why* they differed is less clear. Was it because they were taught differently? Did they differ because they work in different schools where certain norms have developed? As part of the member checking process, I asked each teacher to respond to two sets of questions that are also included in the MSTA survey (discussed in Chapter 6). Their responses are summarized in Table 5.9. The portrait that emerges from all seven teachers is that they are individual *bricoleurs* without dominate super ordinate systemic guidance in their development of classroom assessment systems. The information in Table 5.9 suggests that they also draw from a variety of intellectual as well as material sources in constructing their individual assessment approaches.

Table 5.9 - Case teachers' ratings (1-5) of influence on their assessment approaches

	- - - Teacher - - -						
Rating of influence of	**PB**	**FC**	**BD**	**JH**	**CC**	**BR**	**VJ**
Pre-service training	2	2	5	2	5	3	4
In-service PD	2	3	3	3	3	4	3
Colleagues approaches	4	2	2	3	3	2	4
First teaching experiences	4	3	3	3	4	3	4
District/Science center recommendations	4	4	4	2	5	4	3
Textbooks	3	4	2	4	2	3	3
Trial and error	5	5	4	4	2	3	4

Legend: 5 = highly influential, 1= no influence.

While these responses do not show any dominant pattern or support the division of this group of seven into clear categories, they do illustrate the diversity of influences that these teachers draw upon when developing their individualized approaches to classroom assessment. Although the option existed to rank any of the sources shown in Table 5.9 as having no influence, none of these teachers ranked any of them that way. As varied as these different reports of influences on the teachers' development of classroom assessment practices are, they may not be as varied as the individual teachers' backgrounds. The study does not have much additional data to bring to bear on questions of why teachers assess the way they do in their classrooms. To further understand these teachers, I next turn to a short analysis of their perception of the historical period that they are in and their attitudes regarding assessment.

Analysis of Beliefs about Historical Period and Assessment

In this analysis, I want to explore how these teachers perceive the current historical period and the relationship between the high-stakes accountability and testing in the form of NCLB and their own work in classrooms. This analysis will shed some light on the thinking of these teachers regarding change and systemic pressures. It is based upon their responses to three specific questions (RQ11-13) and their extemporaneous responses to other questions. While most teachers acknowledged the current period as one of important historical change, where new technology is altering aspects of their practice, there were diverse opinions regarding accountability and NCLB.

Of the seven teachers, Paul Bond represented the most insular point of view regarding systemic change. He reported a high degree of stability for his classroom practice. The last decade's changes have not had much impact on his assessment practices, as he describes in Transcript 5.8.

Transcript 5.8 - Paul Bond discusses NCLB and the decade of change.

1) PP: How is No Child Left Behind affecting your job?
2) PB: No Child Left Behind? No Child Left Behind is just kind of like a slogan that that we all just laugh about.
3) PP: laughing
4) PB: Seriously, it like it means nothing to us. It is like something the Bush administration or whatever threw at us with their new logo No Child Left Behind...whatever.
5) PP: Yeah..Um how would you describe the decade of 1996 - 2006 in terms of change of assessment practices?
6) PB: Truthfully, I don't think it's really changed much at all. They change, they monkey with the MEAPS and whatever but basically it's the same stuff.

Paul did refer to kids learning about Pluto from the Internet, so it seemed that he has kept pace with technological change even while his classroom approach remained consistent. He also indicated that some students were difficult to reach and that middle school is the time when many of those students give up. Often, he said, there are home/family issues with these kids, and he expressed a frustration that it is not possible to reach 100% of the kids. He said it is rare to "see parents of the kids I fail at parent teacher conferences." He contrasted public and Catholic schools in terms of parental commitment. However, he also said that "sometimes they [students] surprise you" by reversing the downward trend and doing well. Paul can be said to be a teacher who is both caring and operates within a framework that accepts some degree of student failure as inevitable, although his inclusion of parent communication sheets in his approach attempts to provide an alternative.

Faith Churchill is another teacher who was fairly stable in terms of her practice. Responding to questions about this period of time, Faith described the changes as dramatic, especially changes due to technology. She did not, however, show that her classroom practice was affected or changed in any significant way. While she did express some frustration over the bureaucratic aspects of NCLB, she also gave examples of why accountability measures were needed and expressed no hostility to the MEAP or testing in general. Faith described the current period of testing and NCLB as specifically onerous in terms of its *bureaucratic requirement*s. At the same time, she stated her understanding of the need for NCLB by describing how some teachers and other districts set the bar too low for their students. She said she reviews the annual MEAP results both with colleagues and with the students. When the eighth grade MEAP results showed weaknesses, she told her seventh graders that they all collectively (including her) have work to do. She said she told her class "*we* are deficient" in an area (i.e., interpreting graphs) and that "*we* have work to do."

In Avon Falls, Betsy Dearing has recently experienced budget shortfalls and reductions in the funding available for schools. Sometimes she had to spend her own money on supplies and has revised tests, changing the types of tasks she gives students in order to save on copy paper costs. Given that she was in her seventh year of teaching, her perspective on the changes in education was influenced by her past experiences with the budget and was also related to her development as a teacher who was realizing she could not do it all. Betsy said the current accountability programs such as NCLB are indeed affecting her job. Because NCLB emphasizes writing, English Language Arts had become a focus of the schools; but as a science teacher, she believed that learning content and scientific methods were both just as important and getting the short shrift. She wondered whether, by middle school, students should be expected to already know the fundamentals of language and math and whether accountability assessments could help ensure that they do.

Betsy's co-worker Jim stated that "government mandates" tend to "turn teachers off a bit." And he presented testing in mostly negative terms. He stated he believed it "twists the concept of education" and that the pressures to cover specific benchmarks are taking away from teaching other important topics. When asked about what he expects for the next five years, he discussed his own plans for developing more conceptual types of assessments. He did not comment on how NCLB or some similar accountability system might impact his work.

The most positive teacher about accountability, and the one who was using the most technology, was Christy Connolly. She said that her job would be difficult to do if she had to return to the paper tools available ten years ago. She also indicated, as did some others in the study, that the role of testing was only likely to increase. Christy was the only teacher in the study who strongly supported testing and acknowledged that her responsibilities as a teacher include preparing kids for tests (see Transcript 5.6 turn 10). When asked about NCLB, she responded positively, as Transcript 5.9 shows.

Transcript 5.9 - Christy Connolly discusses NCLB.

1) PP: How is No Child Left Behind affecting your work?

2) CC: I have to say it really helps me stay focused on the kids.

3) PP: Really?

4) CC: Absolutely. Um and I think it's because philosophically I believe that every single student should be successful in my class and so y'know even before they came along with it and when it came along I said you mean some people aren't thinking that way? You mean your not using the test to see how well you're teaching and how well they're learning?
I guess because was an accountant I ... just feel that accountability and always have ... My problem, I don't know .. It is probably not a very popular thing to say but I don't care. I have been waiting for it to get to science.

5) PP: yeah and --

6) CC: I think the problem .. the only problem with that is that the schools are spending so much time making sure the kids read. Gee **what's wrong with that** .. but they do It would be nice if they were doing more content reading and making sure that they know how to do math that a lot of them aren't doing science .. um that that is a shame .. but a lot of them didn't do science before and now they are pointing to it [NCLB] and saying we can't do science but y'know **you never did science before**.

The teacher whose position was perhaps the most balanced was Ben Raminskis. Ben did not seem a fan of high stakes testing, and he described some of the downsides to it. However, each of the three times the subject came up, he volunteered a complementary perspective that there were both benefits and costs to assessment, as Transcript 5.10 illustrates.

Transcript 5.10 - Ben Raminskis discusses NCLB and accountability.

1) PP: How is NCLB affecting your work in science education
2) BR: No Child?
3) PP: Yeah, No Child Left Behind.
4) BR: Well I think probably sticking more to the formal state objectives than I did in the past.
   Um less ... um ... you know you used to do, I used to do units that were more ... I don't know what to say.. things that were almost more extensions than opposed to sticking to this is what should be taught.
   So I think in some ways it's probably ... made it less interesting for the kids in my judgment and less interesting for me.
   But in other ways it's y'know somebody said this is important, the kids need to know it and I'm not teaching it, so maybe I better change what I'm doing a little bit.
   So it's been good and bad.

With Valerie Jones and the Hardy teachers, the questions about how testing was affecting their jobs were not productive. Perhaps this was because their school was in the middle of a restructuring effort that is discussed later in Chapter 7. They all expressed an interest in better performance on the state test and discussed that they had been reviewing annual MEAP results for years.

Across this group of teachers I can identify two opposite perspectives. One extreme is represented by Paul Bond, whose teaching style was regular, utilized frequent tasks, and was little affected by changes in technology and accountability requirements. Christy Connolly represents another extreme, as she was actively incorporating new technologies in her classroom, used an improvisational teaching assessments approach with irregularly spaced assessments, and was actively responding to accountability requirements. Further, as a participant in the development of the annual test, her classroom work was linked to the state's accountability apparatus.

This analysis has not produced pure categories or unproblematic distinctions. Between the examples of Paul and Christy are a range of other beliefs and practices that do not fit neatly into categories or along a continuum. In Chapters 6, 7, and 8 these differences will be explored from other perspectives that, while still not purifying the conceptual categories for seventh grade science classroom assessment from a systemic perspective, will hopefully provide a more complete description.

Seven Teachers, Seven Systems

This chapter began with a discussion of the no-gum zone – the way that one teacher structured her classroom as a different kind of place for her seventh grade students. In using this example, I was both making statements about the nature of seventh graders in schools and inserting a perspective of schools as places that can be thought of as having zones. From an organizational view that encompasses districts and state governments, teachers in classrooms often appear either as units in just one level of a multi-layered system, as many studies in Chapter 2 present them. And it seems natural that in a systemic account, the more that any one component can be simplified based on common characteristics, the more straightforward the overall depiction can be, as additional layers and elements are added to the picture.

The teachers presented in this chapter complicate this simplification, however. They show how their approaches to student assessment are individualized and woven into the fabric of their personal classroom instruction. Even just focusing on classroom work, as this chapter has, it seems difficult to characterize these seven teachers as a unitary group. In their daily work, these teachers use assessment tasks for multiple communicative purposes in what Scollon (2001) might call the nexus of practice. While Scollon focused finely on individual communicative acts, in this study, we see teachers engaging in a conversation with their students that occurs over longer timescales (Lemke, 2000), but with similar features of multi-determination. As the aperture of this study expands in the next chapters to include more organizational elements and more participants, the teachers will for a time become less detailed and more like nodes in a network, before some emerge with detail restored.

Chapter 6

The Case of the Michigan Science Education Professional Association

"I don't want to join a club that will accept me as a member."
~ Julius Henry "Groucho" Marx

Introduction

This chapter introduces dimensions of community and organizational membership into this analysis. As the quote above reminds us, membership in a particular organization is a deliberate act of affiliation. The choice to affiliate may be, and often is, parallel to an individual being included in a conceptual analytic category. For example, members of the Veterans of Foreign Wars may be politically conservative; members of The Wildlife Conservancy are likely to hold pro-environmental views; the Parent-Teacher Association is likely comprised of parents who are inclined to participate in the education of their children. However, as these examples illustrate, simply being in a conceptual group and joining an organization chartered to support that group are different things. When an individual chooses to participate in an organization, he makes a statement about his identity. And, in many cases, members of these associations have access to resources and discourses less easily available to non-members.

This chapter focuses primarily on one organization: the Michigan Science Teachers Association (MSTA). There are three professional associations that are used in this study, and MSTA is the most important for this dissertation. It is the primary science organization in Michigan and is broadly affiliated with other science education organizations as well as the state government. Four of the teachers who were the focus of the last chapter are members of MSTA. It is an organization that is both historically established, having celebrated its 55^{th} anniversary during the year of this study, and connected to many of this study's participants.

I will use two primary types of data in this chapter. The first source of data comes from a survey that I conducted with MSTA members. The survey is used initially as a sounding board for some of the themes raised in the analysis of the seven teachers' practices from Chapter 5. It helps develop an understanding of how common (or unusual) the case study teachers' practices were. It helps when we ask whether these teachers are unique in some ways or if they can be found in harmonic patterns of larger groups of Michigan science educators. While this is a study of the particular, understanding ways to extend the particular examples to larger domains is an important task in educational research, and the survey helps in this extension. The survey will also serve as a resource in the later chapters. It aids in understanding patterns of practice across the state, and from time to time as the dissertation progresses it will be helpful to review what the survey shows in concert with other evidence.

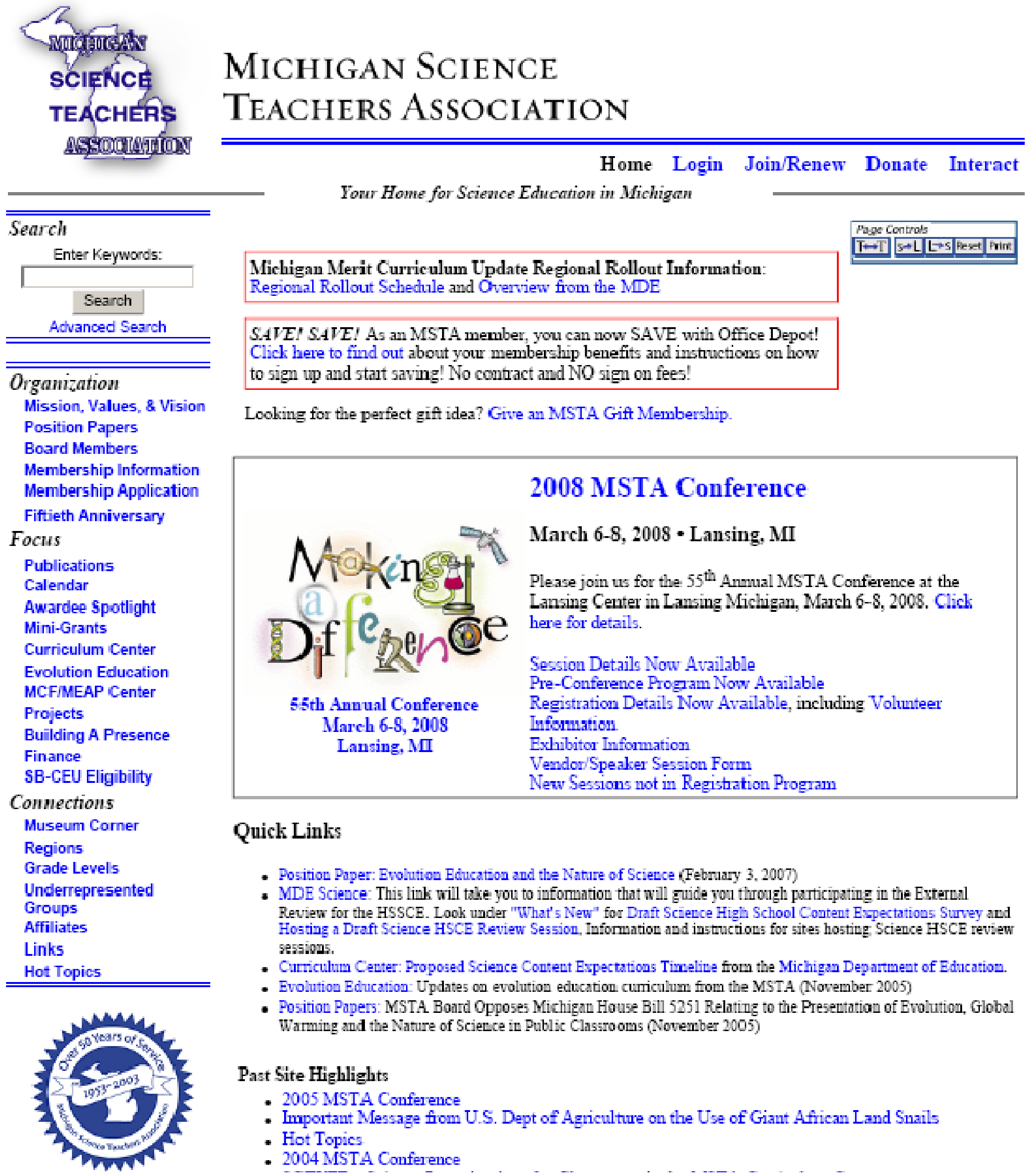
The second type of data source this chapter introduces comes from two individual case study participants. Like the seven teachers profiled in Chapter 5, these two are science educators in Michigan. Their inclusion in the study is not because of their typicality. Rather, the unique perspectives they bring based on the roles they played and their long associations with the Michigan science educational community are what make them valuable contributors. One was president of the MSTA during this study, and the other had been the director of one of Michigan's most influential Math/Science Centers, as well as a long-time MSTA member. These individual cases serve several purposes. They speak to larger patterns and trends in Michigan science education and its support structures in ways the case study teachers and MSTA survey cannot. They introduce into the study individuals with multiple organizational affiliations, a dimension that will play a greater role in later chapters. Through their perspectives, these leaders also shed light on a zone of activities operating between the schools and the state. Michigan has an often irregular and overlapping network of educational organizations (e.g., districts, Math/Science Centers, ISDs), and these MSTA leaders help us to see into that analytical tangle and identify some of its more important relationships and trends to help address some questions of this study.

One driving question for this research involves the prevalence of science education systems in Michigan. In order to properly sketch the context of this study and help understand the historical progression in this state, it is important to understand what alternatives to individual teacher assessment practices and the annual state MEAP test existed at the time of this study. In none of the case study schools was any alternative assessment system found. I bring both the personal accounts of the science education leader cases and the MSTA survey to bear on this question.

Before transitioning to the next chapter, where the systemic aperture is broadened further, I discuss the ways that the MSTA case and these related individual cases allow us to conceive of the relationships between individuals and professional associations more broadly. The MSTA case and the cases of its two leaders provide an excellent example of the principles of Actor Network Theory (ANT) and how this study can contribute to that theory. Within Michigan's complex educational enterprise, as with other complex, multifaceted organizational structures, the relationships between individuals and organizations and between different organizations are often as complex and interesting as the relationships that many people maintain with each other. Organizations can be simultaneously context, meeting place, and mediator, as the MSTA and the two individual cases related to it enable us to see through very particular accounts.

The Case of the MSTA

The MSTA was formed in 1953. At the time of this study it had over 3000 members; published both a newsletter and a biannual journal, and since 1954 had held an annual conference where teachers present strategies and vendors exhibit various products and/or equipment. The Michigan Department of Education (MDE) frequently attends these meetings to make presentations. A review of recent conference programs show the presentations have covered the process for developing new science standards and plans for the MEAP or related state accountability programs. A review of MSTA journals for the past eight years shows that MDE science specialists have also written about these topics for the MSTA membership and presented at MSTA meetings. Frequent notices about opportunities to participate in MDE development of standards and tests were also found in these forums and the Listserv.



MICHIGAN SCIENCE TEACHERS ASSOCIATION

Home Login Join/Renew Donate Interact

Your Home for Science Education in Michigan

Search

Enter Keywords:

Search

Advanced Search

Organization

Mission, Values, & Vision
Position Papers
Board Members
Membership Information
Membership Application
Fiftieth Anniversary

Focus

Publications
Calendar
Awardee Spotlight
Mini-Grants
Curriculum Center
Evolution Education
MCF/MEAP Center
Projects
Building A Presence
Finance
SB-CEU Eligibility

Connections

Museum Corner
Regions
Grade Levels
Underrepresented Groups
Affiliates
Links
Hot Topics

Page Controls

Michigan Merit Curriculum Update Regional Rollout Information:
Regional Rollout Schedule and Overview from the MDE

SAVE! SAVE! As an MSTA member, you can now SAVE with Office Depot! Click here to find out about your membership benefits and instructions on how to sign up and start saving! No contract and NO sign on fees!

Looking for the perfect gift idea? Give an MSTA Gift Membership.

2008 MSTA Conference

March 6-8, 2008 • Lansing, MI

Please join us for the 55th Annual MSTA Conference at the Lansing Center in Lansing Michigan, March 6-8, 2008. Click here for details.

Session Details Now Available
Pre-Conference Program Now Available
Registration Details Now Available, including Volunteer Information
Exhibitor Information
Vendor/Speaker Session Form
New Sessions not in Registration Program

Making a Difference

55th Annual Conference
March 6-8, 2008
Lansing, MI

Quick Links

- Position Paper: Evolution Education and the Nature of Science (February 3, 2007)
- MDE Science: This link will take you to information that will guide you through participating in the External Review for the HSSCE. Look under "What's New" for Draft Science High School Content Expectations Survey and Hosting a Draft Science HSCE Review Session, Information and instructions for sites hosting Science HSCE review sessions.
- Curriculum Center: Proposed Science Content Expectations Timeline from the Michigan Department of Education.
- Evolution Education: Updates on evolution education curriculum from the MSTA (November 2005)
- Position Papers: MSTA Board Opposes Michigan House Bill 5251 Relating to the Presentation of Evolution, Global Warming and the Nature of Science in Public Classrooms (November 2005)

Past Site Highlights

- 2005 MSTA Conference
- Important Message from U.S. Dept of Agriculture on the Use of Giant African Land Snails
- Hot Topics
- 2004 MSTA Conference

Document 6.1 - MSTA Web Page

MSTA also has local affiliate organizations, such as the Metropolitan Detroit Science Teachers Association and the Michigan Alliance for Environmental and Outdoor Education, as well as national affiliates, including the National Science Teachers Association (NSTA) and the National Science Education Leadership Association (NSLA). As the organization's web page (Document 6.1) illustrates, the MSTA is active in state policy processes related to science education, including developing position papers and providing testimony regarding state curriculum and assessment issues.

Because MSTA is the most prominent organization for science practitioners in Michigan, I make it the basis of a case that also reflects on these affiliate organizations. I build the case using data about the membership that comes from public documents, including newsletters and position papers; the survey I conducted, and also the two association leader cases.

As a concrete entity with specific membership, MSTA provides a solid foundation for expanding the depiction of practice that the seven teacher cases provided. This expansion need not target a single category, such as a *typical* seventh grade science teacher. Rather, it looks at teachers of other grades, different science subjects or different types of students, and a broader set of students. Through its 14 regions, MSTA also allows the issue of locality to be discussed, thereby providing other context not represented by the seven teachers to be explored. The MSTA's archive of publications and meeting agendas also provides resources for looking at the evolution of topics over time within the science education community in Michigan.

The MSTA also provides access to other individuals who, as a result of participating in many MSTA affiliate groups and with the MDE in various roles, have a greater awareness of larger issues across regions and over time. I introduce two senior MSTA members, Pete Darmond and Corrine Eaton. Both had been teachers before moving into leadership roles in Math/Science Centers (MS Centers[22]). Pete is also part of a large Intermediate School District (ISD). Their work gives them some insight into what is occurring in hundreds of classrooms and organizational processes in this intermediate organizational zone that relates to classroom activity, including professional development and curriculum planning. While each of them has enough experience to make rich in-depth case studies of science education leaders, in the interests of space I present them only briefly, to illustrate connections in the professional community and to further explicate the nature of the network properties of this study's participants.

[22] This abbreviation is the one chosen by the Math/Science Center Network organization.

The MSTA Survey

As a way to explore the potential for using these seven cases for broader inferential purposes, I conducted a survey of MSTA members in the fall of 2007. The response from this informal survey (n=198) included individuals with different roles and working in different contexts, as Figure 6.1 shows. The survey was designed after the study of teacher practices discussed in Chapter 5. It was designed initially to be used as a comparison data set for what was being learned in schools and from other participants, and many of the questions were multiple choice items influenced by the earlier part of the study. Some questions were open-ended, and some provided opportunities for the respondents to select from a predetermined set of choices and add their own selection/comments.

The survey was conducted over the Internet and selectively displayed sections based on respondent characteristics. There were separate sections for: demographics, experiences, roles, general attitudes, classroom teachers, curriculum specialists/department heads, and for those who had worked with the MDE in developing/reviewing assessment items. An emailed invitation went to the membership from the association management company just before the beginning of the school year in September 2007. This first invitation resulted in 92 responses. A second invitation was emailed in October 2007, resulting in the remaining responses. The survey did not request any identifying information, but rather allowed respondents to indicate their region, general characteristics about themselves, and the context in which they work. The survey cannot be used for statistical purposes due to issues with sampling and the respondents' heterogeneity, but it is used in this study as a source for suggestive and comparative evidence.

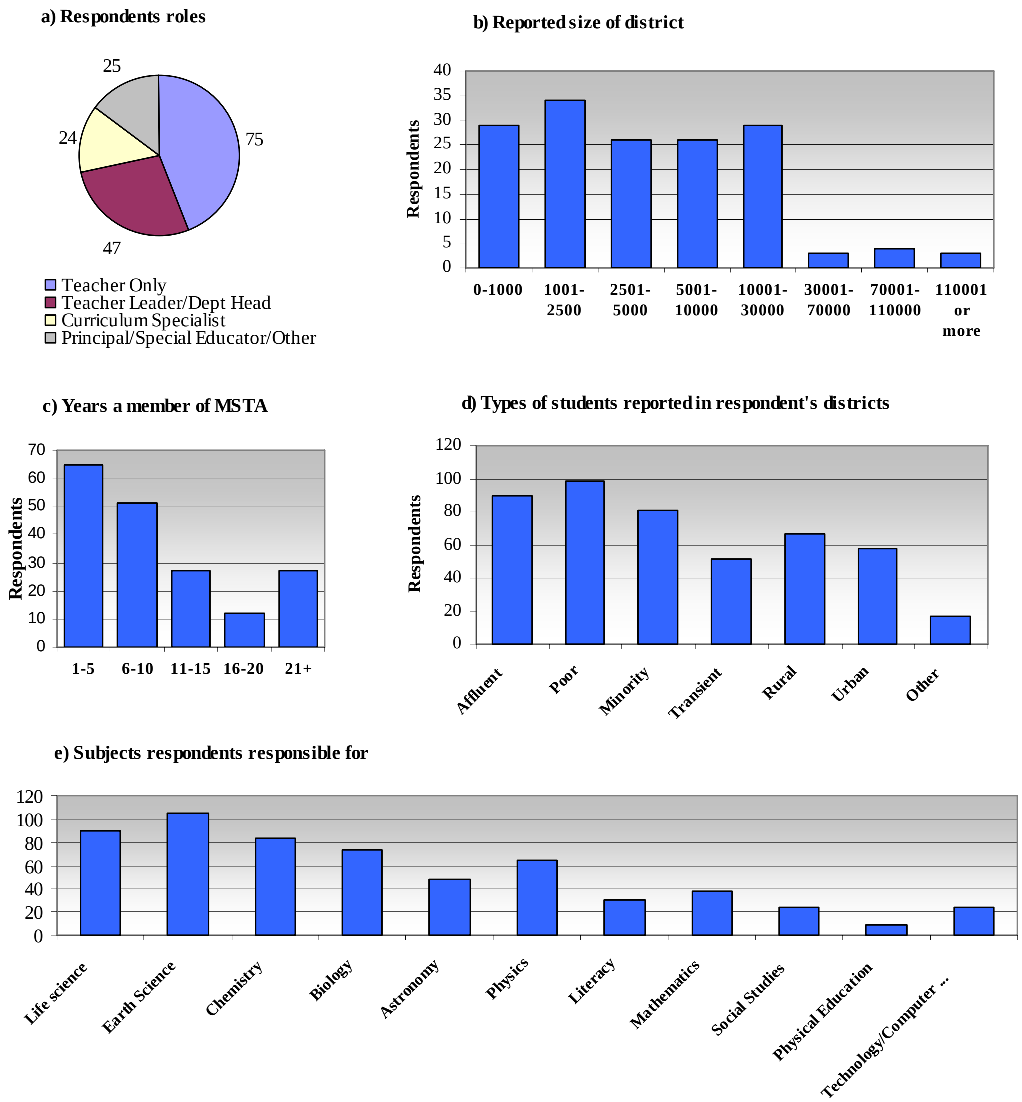


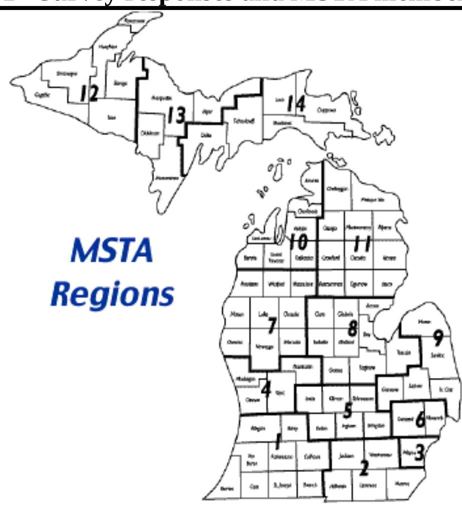
Figure 6.1 - Characteristics of the MSTA survey responses

Determining the relationship between the characteristics in this sample and various education populations in Michigan is problematic, because in many cases the data on the population is not easily available. For example, data on a teacher's years of service is not maintained in any public information system in the state.[23] The state government receives the information from districts, but districts do not always keep records on teachers before they entered the district. However, in some broad ways, the respondents' replies provide information about many important aspects of the state. For example, respondents reported working in different sized districts (Figure 6.1b), having a

[23] The state teacher's retirement system may contain some information that is not specific to the subject or school. My request for information from that system was denied.

range of experience levels with MSTA (Figure 6.1c), and being responsible for different types of students (Figure 6.1d). Furthermore, the survey respondents were 66% women, and of the MSTA members who indicated a gender, 68% are female. In Michigan, 75% of all staff are women.[24] When asked to estimate the number of science teachers in the school who were MSTA members, the response average was about 25%, indicating that the teachers who participated in this survey most likely did not represent the overall population of science educators, but tended to include those involved in MSTA.

Table 6.1 - Survey responses and MSTA membership by division



Region	Survey N	Survey Pct.	MSTA N	MSTA Pct.
1	20	10.2%	294	9.5%
2	9	4.6%	182	5.9%
3	19	9.7%	277	9.0%
4	25	12.8%	414	13.4%
5	15	7.7%	275	8.9%
6	39	19.9%	463	15.0%
7	6	3.1%	79	2.6%
8	19	9.7%	245	8.0%
9	11	5.6%	179	5.8%
10	2	1.0%	48	1.6%
11	1	0.5%	40	1.3%
12	1	0.5%	9	0.3%
13	2	1.0%	24	0.8%
14	3	1.5%	23	0.7%
Other	24	12.2%	527	17.1%
	196		3079	

Cross-case Analysis of Classroom Assessment Activities

This cross-case analysis has the modest goal of further understanding in what ways the patterns exhibited by the seven case study teachers are similar to those of other MSTA teachers. This analysis, which looked at the same topic across different types of cases, reflects a fundamental principle of this dissertation's design that is also used throughout this study. This principle is that different types of cases can be productively compared. The first step in this analysis is to look at the survey's inclusion of different grades; the analysis of the case study teachers and much of the rest of the study are

[24] This is the only comparable figure that I was able to find on the Michigan State website (http://www.michigan.gov/cepi/0,1607,7-113-21423-177922--,00.html). I would anticipate that many of the staff types, such as aides and cafeteria workers, may be more heavily women than men.

focused more narrowly. Although the survey included teachers from many grade levels, the responses suggested that in some respects MSTA members teaching different age ranges employ similar assessment instruments, as Figure 6.2 illustrates. These similarities allowed me to combine respondents from multiple grades for simplification. When I do this, as I do in Table 6.2, I have previously compared the responses by grade and found no appreciable difference.

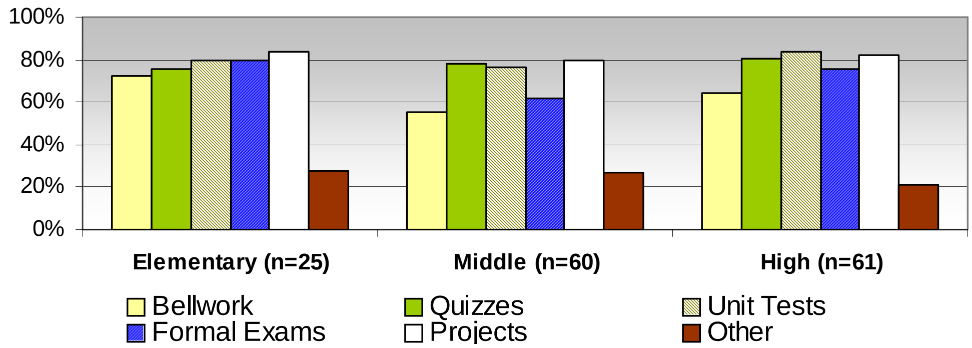


Figure 6.2 - Reported characteristics of teachers' assessment approaches

The survey supported characterizing the seven case study teachers' practices as similar to those of the responding MSTA members. For example, the teachers in the seven cases arranged different types of instruments (quizzes and tests) with often different frequencies of use. Many survey respondents (Table 6.2) reported similar approaches where they arranged some instruments with more frequency than others. The average number of the six categories of instruments chosen in the survey question was 4.7, with no teacher selecting less than three of the assessment types. The case study teachers reported an average of 5.4 assessment instruments. However, this figure included special types of assessments that survey respondents could not include.

When considering the reasons why teachers gave different kinds of assessments, the MSTA survey respondents also showed patterns that harmonized with the case study teachers, as shown in Figure 6.3(a). Like the former, the latter reported that their purposes for giving students assessments were broader than simply measuring knowledge. These reasons included using assessments to reflect on their own practice and building student knowledge, in addition to measurement. Although these responses showed similarities in the attitudes of teachers at different levels of school, they also showed that the teachers did not value all assessments equally. Some purposes (ex: measurement) were highly valued and some (ex: giving students something to do) were valued much less.

Table 6.2 - Reported frequency of different assessment approaches across all teachers

	Bell-work	Quizzes	Unit Tests	Formal Exams	Projects	Other
Daily	50	15	6	0	0	3
Weekly	28	48	12	0	1	3
Bi-weekly	2	34	41	5	0	8
Monthly	2	7	45	54	11	16
Several times a year	2	5	6	41	16	34
Yearly	5	4	5	17	56	42
Total	89	113	115	117	84	106

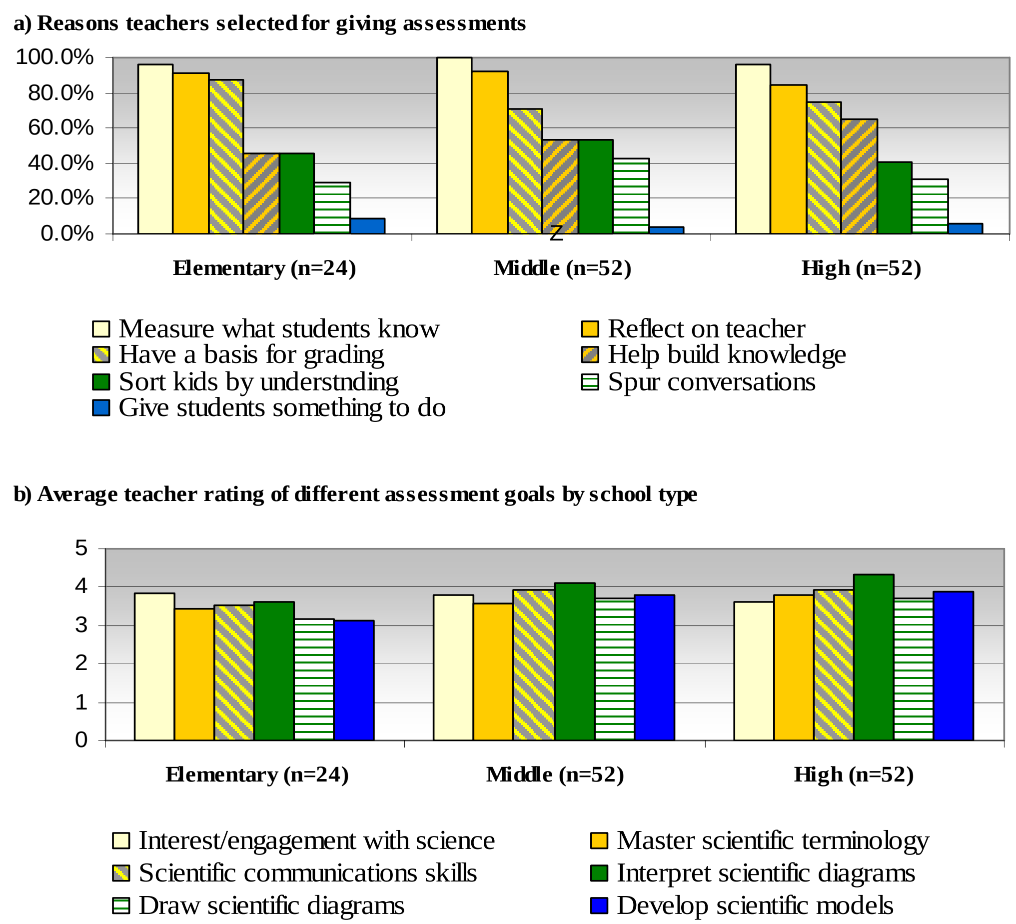


Figure 6.3 - Reasons for and values assigned to assessments by school type

Consistent with the analysis of the case study teachers' assessment practices, survey respondents reported a range of symbolic activities, as shown in Figure 6.3(b). And across the three grade ranges, these different types of assessment activities received similar types of ratings.

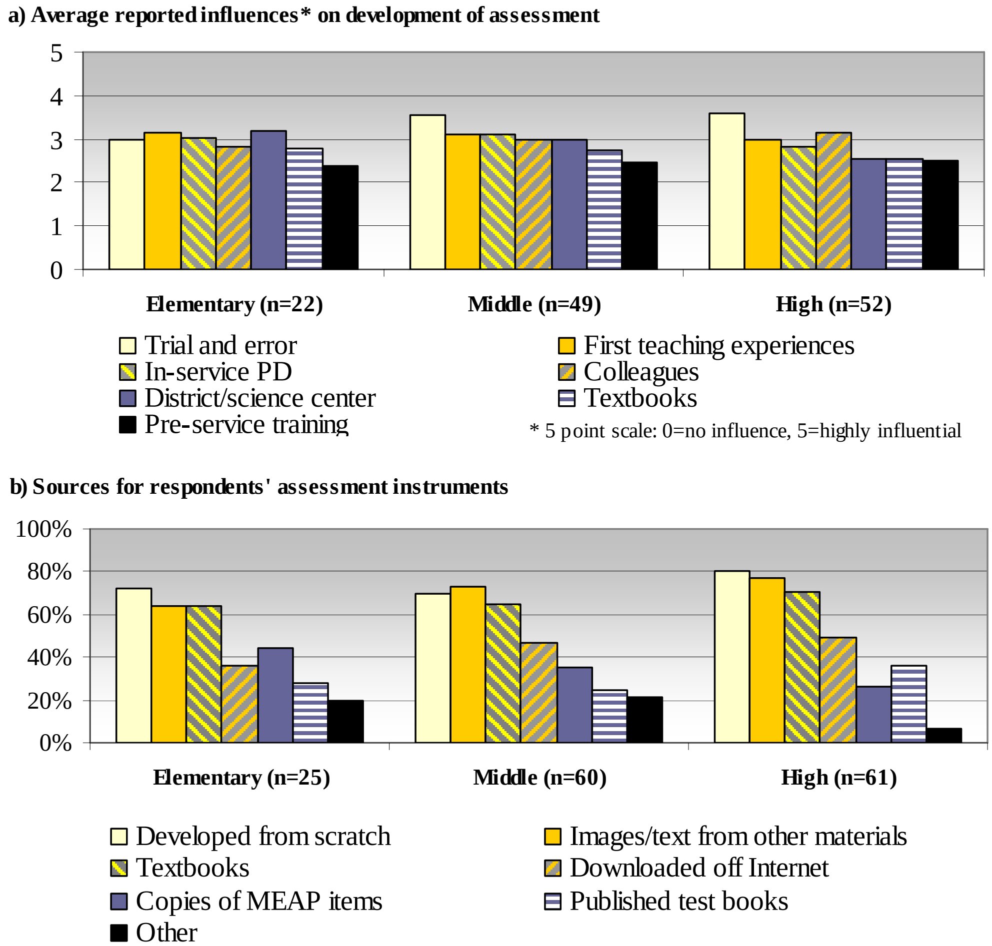


Figure 6.4 - Characteristics of the development of teachers' assessment systems.

When considering the origins of different teachers' assessment systems and the influences that affected how they approached assessment, the case study teachers were not unique. The survey respondents reported a wide range of influences on their assessment approaches (See Figure 6.4.a), similar to the influences reported by the case study teachers summarized in Table 5.9. As Figure 6.4.b shows, many teachers are likely to develop a large portion of their assessment systems themselves, which also parallels the case study teachers.

Through this informal survey, the MSTA case provided a way of comparing specific case study findings with broader perspectives in Michigan. This suggested that in many ways the seven teachers are not unique in their approaches to assessment. A view of classroom assessment practices, illustrated by the seven case teachers, as independent systems of practice that operate without strong uniform influence from either the school districts or publishers is not contradicted by these survey data.

MSTA Leadership Cases

The survey presents MSTA in terms of groups of anonymous individuals responding to structured questions. These responses help to illustrate possible patterns that exist at the state level. I draw upon other parts of the survey throughout the dissertation, but now switch focus from a group to specific individuals. These two MSTA leaders will help illuminate the activities of the networks of support organizations that link schools within Michigan. In both of their accounts, these senior members of the professional community provide a perspective that is not directly related to the seven teachers in this study discussed in Chapter 5. The teachers neither work for them, nor is there any evidence that they know them. Still, their perspectives reinforce the depiction of science teachers working without substantial resources and can help them develop assessment approaches or develop assessment alternatives to the two types of assessment systems common to all schools in Michigan: classroom assessments and the MEAP.

The Case of Pete Darmond

Pete Darmond is currently co-director of an ISD-based MS Center located in McReady County, one of the suburban areas in the southeast of the state. McReady is densely populated compared with much of the state and comparatively rich in resources. McReady's ISD is responsible for roughly 130,000 students in over 20 school districts. The MS Center provides teachers with both educational support and equipment. Pete taught elementary school for 24 years, where he initiated a program that was awarded a ComputerWorld award in 1995 for program innovation and excellence. Web searches show that, since that time, he has either received awards or presented them on behalf of the McReady ISD, where his official title is science consultant. Pete has served as the president of MSTA since 2003 and at the time of this study, was also executive director of the Detroit Metro Science Teachers Association, an MSTA affiliate. As president, he manages the relationships with the association management company and leads the board, which has 50 members that include representatives for each region, school type, and affiliate organization. He has also served on a variety of committees for MDE, including committees on the development of standards and MEAP tests.

Transcript 6.1 - Pete Darmond discusses district commitment to data use and uncertainty about classroom change.

1)	PD: Use of assessment and using data in the classroom on a systemic basis I would have to agree with Vera that there is probably a spectrum. Speaking just for McReady, and I know Phil we had talk about this in a previous conversation that about the Wolsons and um the ISD.
2)	PP: uhhun.
3)	PD: And our ISD had hosted Mark and um his wife I believe Deb .. here at McReady for an an institute .. um for on the Successline what I think its called the Golden Package.
4)	PP: Umhm.
5)	PD: Program that was funded by our ISD and so every every district was invited and I would have to say that we had we had um every district here um plus representatives from the [not named] Public School system here.. who participated .. in that in uh that institute. So there were homes.. there were projects and they were supposed to go back and have projects to create assessments and make that happen. So as far as districts go on the ADMINISTRATIVE level there is there is a commitment to um or there seems to be a commitment anyways um towards assessments .. towards making assessments happen getting the data and uh making the changes the classroom.
6)	PP: umhm
7)	PD: Having said that.. what I don't know is .. from distinct to district to district what their progress is in making that happen in the classroom.

Pete is also a member in an informal group of state MS Centers that he called the “thumb and knuckle group” because of the region (eastern Michigan from Saginaw down to the Southeast Metro area) it is from. This group developed a database of test items for teachers to use, but he was unable to verify who used it or what was done with the information it provided. When asked about alternatives to MEAP and classroom practice for seventh grade science, he could not come up with an example. He did, however, describe a workshop on using data for instructional decisions. The discussion about it in Transcript 6.1 is important because he is mentioning the activities of a software vendor, Successline, and their product called the Golden Package, also known as MI-Tracker (pronounced “my tracker”).

Within this short part of Pete’s discussion, some of the nature of the MS Center and ISD curriculum service work can be seen. The ISD and Center sponsor workshops for educators in their districts, described in Transcript 6.1 (turns 1-3). Other aspects in terms of supporting districts with equipment and curriculum guidance are described in Transcript 6.2 (turns 24-26). In a portion of the interview not included here, Pete described a special science school the ISD runs for students from local districts. This school pays for part of his salary, which provides him a guaranteed position -- something not always the case for other science specialists who are not full-time teachers.

Successline has a strong presence in Michigan and connections to several schools in this study. The company’s Golden Package/MI-Tracker software is a combination of data visualization tools and consulting services oriented towards MEAP results. According to Successline, the package is capable of handling other types of data. But, at the time of the study, they produced versions specific to the annual tests for several states. In all cases, these special versions were tailored to state tests. Successline, which appears again in later analyses, is a good example of how organizations in this study can be linked. Not only do they share a common technology and information architecture, their members also will attend similar professional development events associated with it. The discussion of the Golden Package workshop highlights the often permeable boundaries between ISDs and school districts, since the other school system representatives crossed an ISD boundary to join this workshop.

This example shows how someone in Pete’s position is able to see the organizational agenda and what may be occurring in many classrooms. Not only is McReady a large ISD, it is bordered by other large ISDs in this densely populated part of the state. Because of this, Pete has a view of science classroom instruction much wider than many other members of the MSTA. Pete is also a leader in the Math/Science Center Network. This group is the collection of over 30 centers in Michigan and is a forum where senior science leaders can connect and share information. It is concerned with many of the same issues, but is much more selective than MSTA. Through the MS

Center Network, Pete was also able to speak to the changes that are occurring across the state in terms of funding for science education support systems, as shown in Transcript 6.2. The meeting he is discussing is an annual leadership retreat attended by MS Center directors and MDE science consultants.

Transcript 6.2 - Pete Darmond discusses changes in the Math/Science Center Network

1) PP: If you had to predict what was going to happen with math/science centers, generally, are they going to be around? Are they going to be smaller? Are they going to be larger?

2) PD: Um...

3) PP: Hard to s.

4) PD: Well it's already beginning to happen. At our last meeting we were asked for the purposes of our website..do you have the link to the website at all?

5) PP: I do.

6) PD: We were asked to discuss if there were any NAME changes that had to be indicated on our website and as I understand Stanley/Wushutunesa those two counties worked together to have one one Math/Science center .. and I believe I can verify this I believe the Wushutunesa the um the the

7) PP: They are going to bring it in house?

8) PD: They are gone. They're just gone. And so..

9) PP: Just doesn't exist?

10) PD: Huh?

11) PP: It just doesn't exist?

12) PD: Right. And as the money goes away as the money has gone away as the money is going away at this point um you're finding that there are certain centers that are simply going to go away.

13) PP: Wow ... and um and what will that mean for?

14) PD: Well it means as far as support for math-science whatever ISD or ISDs tended to be in that area would have to pick up the support... they would have to then uh pick up the support they'd have to pick up the support role.

15) PP: And it' not Wushutunesa doesn't have money. We know they have money.

16) PD: Right.

17) PP: We know they have a lot of money.

18) PD: I have to verify that too..but I know at the last one.. basically the name change because of what was happening there due to layoffs or due to contracts not being renewed or maybe a person retired and not rehiring somebody those things are going away. At the last conference when I looked at the faces at the last Math Science Center Network there were a lot of new faces and many of them were not even science and math people believe it or not.

19) PP: Really?

20) PD: Nope. They are not even science or math people. The ISD as the fiscal agent receives the uh receives grant money that comes in whatever amount that might be. And y'know if they've had to make cuts in personnel but they still receive the the grant money of course they have to send an individual or designate an individual to attend the meeting. And in some cases you have .. the person representing the science math center network not even being a math science .. uh content specialist or person, they might be an ELA language arts person.

Transcript 6.2 - Pete Darmond discusses changes in the Math/Science Center Network

21) PP: And that's not encouraging at all.

22) PD: Yeah oh yeah and that's the reality we're dealing with. The fact that the money's going away the Math Science Center Network itself has had to look for other funding sources .. so they've been in dialog and working with the Kellogg Foundation

23) PP: Right.

24) PD: Uh get to basically to get assistance in how to go about to get our message out there a lot of for the most part uh Philip even with the funding cuts it is a two edged sword most of us roll up our shirt sleeves and off we go and making things happen despite the loss.

And it's a two edged sword because you could say oh if we say oh look at a place like let's say McReady we are offering this PD and whatever and its true its not people. My salary and Mitch's and Cathy's we're not dependent on those math science center dollars because we're hard ... we're part of the ISD uh staff.. but as far as what we can get and offer in terms of resources and equipment .. and other kinds of things .. that's going away too but we're trying to figure out how we can make that happen y'know with um less dollars.

25) PP: Right.

26) PD: And so .. if you if you still make things happen with fewer dollars the two edged sword is you can say they are still making things happen and we can cut them further .. that's the fear y'know .. But on the other hand everyone is trying to make things happen despite having the fewer dollars.

In this short excerpt from Pete's interviews, the highly interconnected nature of this study's evidence and also of professionals in this area of the system is shown in several ways. The MS Center run by Stanley and Wushutunesa counties would have served three of the seven case study teachers in this study: Faith Churchill, Christy Connolly, and Ben Raminskis. Because their schools are in districts located in Wushutunesa County, an active MS Center may have been raised in their discussions of professional development or assessments. However, none reported any involvement with the center. This same center will appear again in discussions in the next chapter in a way that showcases the multiple perspectives this study design supports.

The Case of Corrine Eaton

Corrine Eaton is the director of the Challenge River Math/Science Center located in rural and central Michigan. She provides another complementary perspective from a senior member of the science education community, and she speaks to other systemic aspects of science education in Michigan and assessments practices within it. In response to my questions, she stated that she had "been an MSTA member for years. I was on the executive board for 11 years and received the MSTA Distinguished Service Award in 2005." She indicated the center's staff have presented at MSTA meetings for the past 17 years and were on the team writing the newly adopted K-7 science content expectations. She has also served on the state's Content Review Committee for MEAP (discussed in Chapter 8) and served on several other MEAP committees over the years.

Challenge River occupies a former school building, where it has a staff of math and science experts and includes a special math/science/technology program for high school students, who take some of their coursework in the center and the remainder of their courses in local schools. In Transcript 6.3, Corrine discusses the local science expertise in her region and raises the issue of disseminating student performance information through a system.

Transcript 6.3 - Corrine Eaton discusses the science curriculum support in her area.

1) CE: Well we only have, in our three county area we have no curriculum specialist in science..
Challenge River has one in math and she's the math and science coordinator and so she relies heavily on us to assist her.

Other than that there isn't a curriculum specialist in our three counties that has a science background And there at the ISD there isn't a science consultant. The Data Director person has was a science consultant but has moved into the data area so there's really not even in our Intermediate School District with a science consultant any longer. So

2) PP: So you're it

3) CE: We're it. And I keep telling the school districts. And I can't be the science curriculum director for every single school district K-12.

And so we've done a lot with building science leadership in our area. We meet with science leaders. We have a leader in every building. And that is part of this program as well as we identify leaders we try to build the capacity in the buildings. But we y'know can't be the person who helps with all the textbook selections and y'know does all the pacing guides and orders all their equipment. We just can't and it does show. Decisions are made in school districts without um all the information they should be because no one knows all the information to look for or ask.

The Challenge River center recently became involved in supporting science education in its local districts' middle schools, as Corrine describes in Transcript 6.4. This gives her a window on what is occurring in the schools that contain seventh grade teachers, which this study is concerned with. This example is also related to Pete's comments in Transcript 6.2 (turn 22) where he discussed the Kellogg Foundation and their support for the MS Centers.

Transcript 6.4 - Corrine Eaton discusses her centers work in middle schools.	
1)	CE: I've gotta tell you one more thing. It's [the new middle school initiative] really not all written up like it should be because if we have fifty extra hours, we spend it in the schools.
2)	PP: Right
3)	CE: And so it's nothing we've ever advertised or submitted. You know we have a pretty good logic model now for the current grant because W.K. Kellogg Foundation came to us and said all right you did this in the elementaries.. I asked them for some marketing money for our center and they said we'll give that to you IF you take on the middle schools
4)	PP: uhhuh
5)	CE: And so I did and they have provided us some resources and so that's why I was able to hire a new person.

Challenge River is one of the most influential of the 33 MS Centers, since it operates a science kit operation reputed to be supplying about a quarter of the state's elementary (K-6) schools.[25] The elementary schools that feed Paul Bond's Swallow Middle School and that feed Betsy Dearing and Jim Heinrich's Avon Falls Middle School use these kits. According to the coordinator of the district science center that manages the kit program, Valerie Jones at Hardy Middle School, reported to me that the program was modeled after the highly successful Challenge River program.

Through her leadership of the center's kit program, her work with middle schools, and her service to the state in terms of the development of state standards and assessments over the years, Corrine has a broad perspective to apply to one of the important questions in this study: how prevalent are science assessment systems for the seventh grade? However, when asked this question directly, Corrine was unable to come up with a single example. Upon hearing of my interest in seventh grade assessments, she declared, "That's the BLACK HOLE … It's our biggest issue: seventh and eighth, but you know it starts in seventh and so if we could fix seventh, we've got it made..." The conversation with Corrine then led to data warehouse efforts that she was aware of, but to no reports of science assessment systems.

[25] Corrine provided figures as high as 27% several times. I repeated this to several informed participants, including Pete Darmond, who all supported that assertion.

Cross-case Analysis of the Question of Assessment Systems

In introducing this professional association case, I have highlighted two kinds of evidence. The first is the anonymous survey that illustrates how some characteristics of individual teacher practice (as described by the seven case study teachers) might be related to that of other teachers elsewhere in the state, and possibly in other grades. The second is personal accounts from professional community leaders, who described broad constraints on resources and expertise for science education along with some ways educational organizations can be linked, including through MS Centers, information technology, and instructional materials. I want to bring both types of evidence to bear on one of the key questions for this research: the existence of viable science assessment systems that could be used to supplement the annual MEAP and the systems that teachers develop for their own classrooms. Both Corrine and Pete portrayed seventh grade assessment as largely unaddressed in systemic assessment efforts. The survey also included a question on this topic, asking respondents whether their school used an assessment system to supplement the MEAP and classroom assessments, and if it did use such a system, what it was used for. The responses summarized in Figure 6.5 indicate about a quarter of middle school respondents named an assessment tool or product that was used in addition to the MEAP.

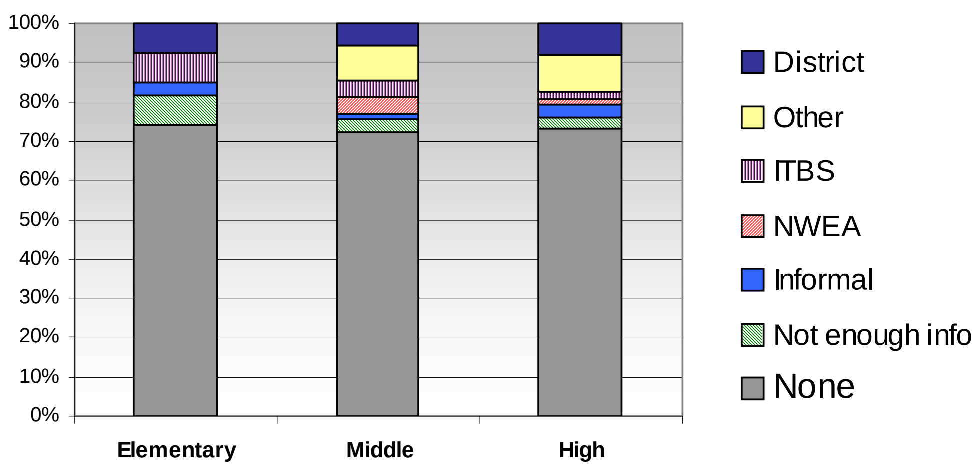


Figure 6.5 - Percentage of respondents' reporting use of assessment systems to supplement the MEAP and classroom assessments

The personal accounts and the survey responses do not seem to agree. To probe this discrepancy, I discuss the nature of the survey respondents who indicated they used a system.

Many of the named testing systems were standardized assessment products (ITBS, NWEA, and other), while some were unspecified district or local assessments. Not one respondent, however, provided a detailed description of what the assessment system was used for. The few responses that stated any purpose at all gave general descriptions such as "pre-assessments gauge student understanding" or "to provide uniform curriculum and common standards of student achievement." One possibility for this discrepancy with Pete and Corrine's responses is that these survey responses come from regions other than those that Pete and Corrine are familiar with. But as it turns out, about half of the respondents indicating they use a supplemental system came from the same parts of the state that Pete and Corrine's own Math/Science Centers are in, leaving this explanation unsatisfactory. Since the survey only asked one question related to these alternatives, there is little additional evidence within the survey to analyze this discrepancy further at this time. The explanation that respondents provided answers indicating a greater use of information than exists in practice is consistent with findings that practitioners over-report their use of data as found by Ikemoto and Marsh (2007), discussed in Chapter 2. The survey question was not as comprehensive as would be needed to answer this question more completely. The important question about how widespread other assessment practices may be for middle school science remains. I will take it up again later in this dissertation from an entirely different perspective.

Reflection on Relationships

The MSTA case and the two MSTA leader cases summarized in this chapter illustrate two important principles about how Michigan's science education organizations and individuals are related and a particular example of more general principles about ways that individuals and organizations can relate.

The relationships of the seven case teachers to MSTA vary. Four of them (Faith Churchill, Christy Connolly, Betsy Dearing, and Valerie Jones) are active members and two (Jim Henrich and Paul Bond) are not. Ben Raminskis used to be a member, but is no longer. Within just the study's seven teacher cases are examples of variable participation between individuals and groups. Appreciating the contrasts that can exist in the community membership is an important step in trying to extrapolate what we see in these seven cases to the broader state context. It is important to understand the evidentiary weight that could be applied to claims made by different individuals. These relationships are conceptually illustrated in Figure 6.6, which shows the unknown population of Michigan science teachers and MSTA members, including many current teachers as well as non-teachers such as curriculum specialists, principals, and other professionals who have a connection to science education.

In introducing Pete and Corrine, I brought in the perspective of highly active and connected association super-members. The two leaders profiled in this chapter were not typical, but special. Their tenure and current responsibilities in the community allow them to comment broadly on what is likely occurring in many classrooms, schools, and districts. Their positions allow them to see and comment about important events and trends that they have observed in the community. By virtue of their organizational roles, these super-members were an important resource for explanations that members with less involvement may not be able to provide.

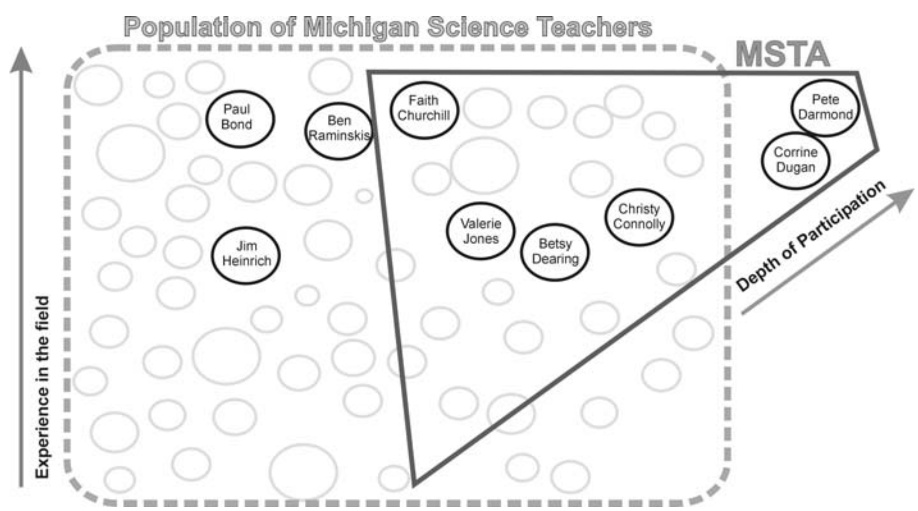


Figure 6.6 - Conceptual schematic of MSTA in relationship to the larger group of science educators in Michigan

As an organization that operates independently of the formal structures of the educational enterprise that include schools, districts, and the state, MSTA's professional community provided opportunities to see those structures through another lens – to appreciate how they are related. While a single organization, MSTA is connected to other organizations including the MDE and its affiliate professional groups (ex: the Math/Science Centers and the local science teacher groups) and thus provides a pathway for understanding the network of organizations in the state.

In looking at how we can consider organizations as related, the MSTA leader cases provided evidence of the non-hierarchical linkages that exist relevant to science education. Through just a few accounts, some particular details of districts and ISDs, districts and MS Centers, funding agencies and M/S Centers, and technology/material providers and different types of educational organizations became visible. Instructional materials, in the case of the science kits produced by Corrine Eaton's center; the funding organizations in the example, or the Kellogg Foundation mentioned by both Pete and Corrine are rarely studied in science education, but they could be of significance to many classrooms as materials are enmeshed into daily practice (Ball & Cohen, 1996; Davis & Krajcik, 2005). Funding agencies, via their investment in strategic programs, may provide similar linkages but at levels removed from the classroom. The data warehouses, including MI-Tracker and Data Director, were also discussed by both leaders, previewing future discussions in this dissertation.

The next chapter will build upon this one by introducing more details about the connections between the case study teachers and both other individuals within this study and various organizations, as well as the connections those organizations have to each other.

Chapter 7 The Evidentiary Web

> Six degrees of separation between us and everyone else on this planet. The President of the United States, a gondolier in Venice, just fill in the names. I find it extremely comforting that we're so close. I also find it like Chinese water torture, that we're so close because you have to find the right six people to make the right connection...
>
> ~ Ouisa Kitteridge in *Six Degrees of Separation* (Guare, 1993)

Introduction

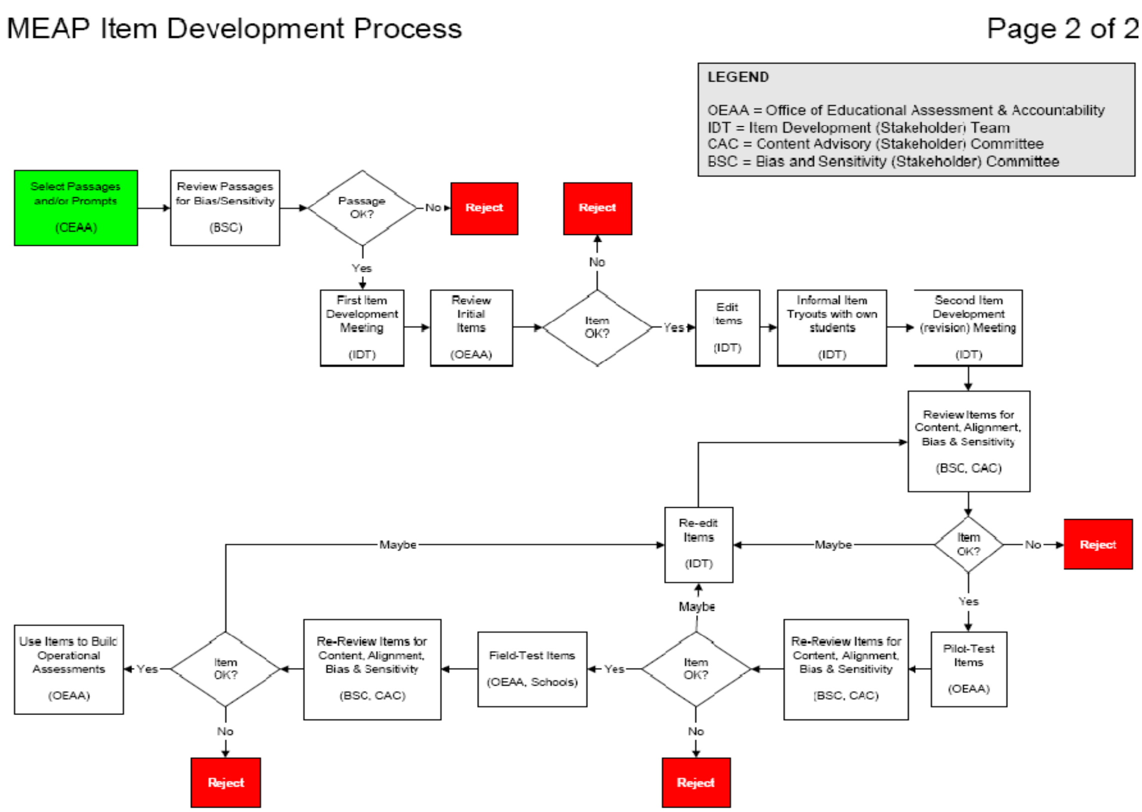
This chapter builds on the preceding two by first exploring issues of organizational context, starting with the five schools that the seven case teachers work within, and then expanding the description of the intermediate levels begun in the last chapter. From the foundational cases of the teachers, I introduce cases for their schools, along with other individual cases for principals, curriculum specialists, and district/ISD staff, to assemble the pieces of the study's evidentiary web from the broader perspective of the state. From Michigan, two important components of this web are added. The state government's MEAP development process and two additional Michigan professional associations form the basis of cases that join the MSTA as representing state-level concerns. As with the connections between the teachers and the MSTA, these new cases introduce relationships that vary by individual. Rather than searching for typical actors, it is the particular connections that make some individuals and these cases relevant. As with Pete Darmond and Corrine Eaton discussed in the previous chapter, some participants' unique positions are used to provide important perspectives on the middle organizational layers and historical trends in this diverse state.

Extending the Case Structure

As I introduce the next set of cases, for each one I will provide a short overview of the individual or organization and describe various connections among the expanding set of cases in this dissertation's data. There are two types of people who are discussed below: case participants, whom I also call *participants*, and non-case participants, whom I also call *informants*. The non-case informants are shown in various graphics, and in some cases their voices are included in the discussion. But the core analyses focus on the case study participants. From this point forward, I will frequently use the terms "participant" and "informant" in this way.

Cases of MEAP Development Process, Joshua Martinique and Roscoe Ellis

The case with the broadest geographical reach in this study is the MEAP test development. This case focuses on those aspects of the development process that bring groups of teachers and professionals together to develop MEAP assessment items within a two-year development cycle.



Document 7.1 - MEAP development process flowchart

Within this formal process, described in Document 7.1, a given test item takes two years from initial drafting by teacher writers to potential incorporation in a bank of test items. This process includes human reviews and student trials that generate statistics that help in understanding possible design flaws in the items. Some of these review activities are discussed in the next chapter. Within this case, I have chosen to focus on two individuals as cases whose roles are integral to this process: Joshua Martinique and Roscoe Ellis.

Joshua Martinique (JM), PhD, is a youthful middle-aged man who had been working in the Office of Assessment and Accountability (OEAA) for only a few years at the time of this study. He had recently been promoted to Manager of Testing from the position of psychometrician. In his management role, he has overall responsibility for the MEAP development and administration process in addition to his more technical responsibilities in psychometrics. His undergraduate degree was in comparative Mayan linguistics, followed by a Master's degree in educational technology. He studied educational statistics in a leading department of measurement and quantitative methods at a large Midwestern university. He actively supported this study and was the person whom I observed most in his professional work, as I attended five conference presentations where he addressed testing professionals. I also met with him several times in the state office building. At one conference, he and a representative of the U.S. Department of Education commented on broad goals and technical issues associated with measuring students over time, with Martinique taking the position that challenged the federal approach to cross-grade calculations for assessing a school's adequate yearly progress (AYP) and the federal representative providing the policy rationale.

Martinique considered himself an outlier among his (national) professional community of testing professionals, because he believes that the concepts being tested are often more complex than the statistical models they use. While the typical test for student achievement is commonly built around a single measurement scale (unidimensional model), Martinique favors the use of multidimensional models where possible. He believes that science education, by its nature, is not well suited to testing using a single measurement target (construct) as he discusses in Transcript 7.1[26].

Transcript 7.1 - Joshua Martinique describes challenges measuring science	
1)	JM: [science] is much harder than the other fields because the constructs change so much..from one grade to the next.
2)	PP: Really?
3)	JM: Science is much more segmented than others.. y'know if mathematics was like science..
4)	PP: hmhm
5)	JM: um we would feel a lot less comfortable with the measurement. Simply because physics is ..the **content** of physics ..the underlying ideas may be relatively similar the scientific method and so on but the underlying ideas but the **content** of physics versus biology versus earth science versus uhm. You have those different pretty discrete sets of content skills that you don't really have in other content areas that makes the measurement of science much more difficult.

Roscoe Ellis (RE) is a tall, large man in his fifties who is the science consultant for the OEAA. He coordinates various subject matter activities for the science assessment development process. Following the lead of his management, he was a supportive participant and provided me with access to various documents and meetings related to the development of the MEAP items. Roscoe has never taught in public schools. Before joining the state government in 1999 he worked in medical education. His undergraduate degree was in psychology, and he has a Master's in business administration. His professional training included statistics, and at one time he considered pursuing a PhD, but decided against it. I observed him with groups of science educators, and on several occasions he showed an understanding for the nuances of science education across different grades. Like many of the people I talked with who worked for the state government, Roscoe seemed to hold a fundamental belief in the value of testing, as Transcript 7.2 shows.

[26] As a matter of discussion, Martinique is actually joining two related issues into one in his discussion. The segmented nature of science, where the same subject area is used for different aspects (e.g., life science, physics, geology) that can often be taught in any order, and the construct change across grades are different. The presentation I saw him make with the representative from the U.S. Department of Education regarding AYP calculations used math as an example. That conference addressed longitudinal measurement, and many speakers addressed the challenges of cross-grade construct mapping using literacy and math examples.

Transcript 7.2 - Roscoe Ellis describes his belief in the value of assessments

1) RE: How the knowledge that you try to apply in science to everyday thinking to the high school graduate...
I drive by on the way to work a... all students are supposed to know certain parts of knowledge according to the state of Michigan. As I'm driving by I'll see the guy laying gas pipe along the side of the road.. standing there by a shovel as he is waiting for the tractor to move in and I sort of tease myself: How did he do on that question on meiosis the other day?
2) PP: Really? [laughter]
3) RE: And when you think about it in reality so all students are supposed to know this...
4) PP: That's funny.
5) RE: And or he's taken .. he becomes a grandpa later on and can he really explain why the sky is blue? You know that's an old reading commercial or something like this.

   And y'know we cover that.. both of those questions

   And maybe as a guy laying pipe, he knows about the soil types, he knows about erosion he knows about the ditch and the water and the drainage etcetera, etcetera or thermal expansions, or I gotta get it below the frost level because of this .. if he is just a high school graduate who had a little technical school and he became a pipe layer for a gas company.

   These are some of the things I am interested in.

Joshua and Roscoe, like many members of the state department of education, have connections to professional communities, including the MSTA. Joshua regularly makes presentations to many professional groups including, on occasion, the MSTA. Because of his role as science consultant, Roscoe interacts more with the MSTA and its affiliate groups. This interaction, he reported, accounts for 20% of his formal job description. Roscoe also presented to the science center leadership meeting discussed by Pete Darmond in Transcript 6.1. My review of the available MSTA journal issues showed one article by him and several by his predecessor describing updates on the state's development of standards and assessments. Roscoe also recruits participants for the MEAP development process from the science education community. He looks for individuals with a range of specific science knowledge so that the committees have broad representation[27] aligned with the science standards.

[27] This was evidenced in observations of the committees and in reports from participants such as Christy Connolly.

In addition to Joshua and Roscoe, I met or spoke with seven other members of the OEAA who are involved in various aspects of testing. Patty Hollander (PH) and Vance Dorn (VD) are part of the group that develops a parallel assessment system for students with cognitive disabilities. Patty was trained to be a middle school science teacher and Vance was originally a school psychologist. Both make presentations to various professional groups and were interviewed in enough depth that a case could have been developed for them. Bobby Black (BB), PhD, is a veteran of state assessment development. Trained in aerospace engineering, he was a math teacher and principal before working in assessments and most recently ran the assessments and evaluation group for a large city school system in the Midwest. I had many discussions with Bobby, who could have been the focus of a case as well, but that case was not developed in the interests of space. I also had a lunch meeting with Don Justice (DJ) who is responsible for the OEAA technology services; Art Gatt (AG) who manages financial aspects of OEAA operations; and Pete Stomber (PS) who handles a variety of issues involving school districts. Outside of OEAA, I met and had several conversations with Keith Roberts (KR), the science consultant in the Office of School Improvement, where he is Roscoe's counterpart in managing the effort to develop new content standards for science (scheduled to be released in 2009). These informants provided useful perspectives that helped me construct the MEAP case.

The Schools and Principals

I now return to the teachers introduced in Chapter 5 to fill in more information about the schools they work in and to discuss some additional related case participants. These new cases will introduce new relationships into what will be sketched later in this chapter as a network of interrelated cases.

This network will have a topology with variable relationships, as some cases are more connected than others. And, much as Chapter 6 illustrated, some individuals have stronger institutional relationships than others. When considering this network topology, it is important to also consider other heterogeneous aspects of the schools included in this study. Not only are the schools different in terms of size, demographics, location, and organizational structure, but their participation in this study also varies.

One area of heterogeneity is the science support structure of the five schools. In the smallest school, Swallow, there is no dedicated curriculum specialist. Rather, the principal of the combined secondary school serves this role for the entire district.[28] Conversely, Avon Falls has a dedicated curriculum specialist in the district who covers all subjects. Hardy's district has a dedicated science person who runs a small math/science center (not part of the MS Center network discussed earlier) to support the district's public and private schools. Dorchester had no curriculum person at all and only a science department head in the high school. And although Crimson's district does have a curriculum department, it had recently eliminated the position of the science specialist. Both Crimson and Dorchester had been served by the Wushutunesa/Stanley MS Center that Pete Darmond indicated had been recently closed. More information about that center is forthcoming in this chapter.

Variation was also evident in the location of science centers in this study. One center, Challenge Creek, is located within a school district, but also houses the substantial materials kit program that serves about a quarter of the state. The one that services Hardy is located within a district, but only serves that district and a few small local schools. The center that Pete Darmond manages is located within a county-based ISD and serves dozens of districts. And the small center that serves Swallow is located in a multi-county ISD where they not only serve a larger geographical area, but do so with less staff and programmatic services.

The Cases of Swallow Middle School and Burt Wainwright

Paul Bond (PB) teaches at Swallow Middle School (Swallow), which is housed in the same building and shares the same principal as the high school in a small rural town located just off a major U.S. highway connecting two medium-size cities in Michigan. Near the town is a large race track, the area's most prominent tourist attraction. Just after turning off the main road, a visitor encounters modest homes and Main Street businesses. The elementary school building is located around the corner from the secondary school in town. Both buildings were well maintained. They seemed well-furnished and spacious.

[28] In the following year, a curriculum specialist was hired for the district, but focused mostly on elementary grades.

After only a short drive, the landscape changes to rural. The school district itself extends into the surrounding countryside and serves a population of around 700 students.[29] The district's population under the age of 18 at the time of this study was 93% Caucasian and 4% Latino.[30] The average income for the town was just over $40,000. Six percent of the students lived below the poverty level.

Transcript 7.3 - Burt Wainwright discusses how assessments take focus.	
1)	PP: In terms of your job as principal, what are the areas where assessments have no impact in terms of what you need to do?
2)	BW: I think they have taken more and MORE and more of my attention, teacher attention, and students' attention than ever before. I don't necessarily see that as a good thing. Um it is what it is. In the social studies world we were always very much feeling that you could not test all of the things you could learn in a social studies class. How can you test compassion? How can you test some of these other things and you CANT. How do you test in a science setting? How do you test ethical uses of technology? How do you do that? But that's got to be part of what you teach as you're going along.. it's got obvious weaknesses but it is a major part of everything we do..and because our reputations are therefore to a certain extent our **existence** as a district is tied to how you do on these assessments. And just to finish that idea science as a case in point some years they will throw on a bunch of stuff having to do with space and other years there's nothing. So how so ..it is random.. you try to just forge ahead and teach what you're supposed to be teaching y'know because you can't just chase that around you're never gonna match up so sometimes the score comes back and everybody looks like they took a big bath on it because something was tested that wasn't expected. The state's gotten much better with that lately by the way much more clear about this is what we're looking for.

Swallow was the first school to participate in the study. I visited there five times and observed one of the faculty meetings where MEAP results were discussed. In addition to Paul Bond, I also created a case for the principal, Burt Wainright (BW). He is a tall man in his early forties. Before moving to Swallow five years earlier to become the

[29] National Center for Educational Statistics 2005 data file.

[30] 2000 census information provided by the National Center for Educational Statistics. This same resource is used for all districts.

new principal, he taught government and civics near Grand Rapids for 16 years. Over my five visits to the school, I saw Burt interact with students and staff. He presented a businesslike countenance that could be interpreted as either parental or administrative. In the staff meeting I observed, he crisply articulated the school's goals and directed the subsequent group work. In his administrative capacity, he shared concerns about the timing and validity of assessments, the needs of special education students versus those more mainstream, and the appropriate interpretation of the school's results by staff members and the community. The appropriate interpretation, in his view, considered internal factors such as which students were included in a given test, and external issues such as how the school's performance compared to schools of similar size across the state. As a school leader and former teacher, Burt presented a balanced view of assessment in Transcript 7.3.

In addition to Burt and Paul, there were three other informants for the Swallow school case who were not used as separate cases themselves: Angie Toliver (AT), the sole math science consultant for the ISD that serves Swallow (compare her singular role with Corrine's staff); Cathy Amazingly (CA), a sixth grade science and social studies teacher, and Toni Donnard (TD), a retiring high school science teacher and the science department head for the district. Swallow also had access to the Moodle course system through a project their ISD had begun, although they had not begun to use it at the time of this study.

<u>The Cases of Avon Falls, Stan Dubovski, Janey Fess, and Bob EnSpania</u>

Avon Falls Middle School (Avon Falls) is in a mostly rural region in the center of the state within an hour of the state capitol. At the time of this study he district had a population of just over 16,000, with 3,300 students in six schools. The mean household income was just under $21,000 per year. The student population was 95% Caucasian, with no other group comprising even a single percent of the remainder. The school system had one high school, one middle school that contained seventh and eighth grades, and an intermediate school in a building connected to the middle school for grades 5 and 6. Three elementary schools fed students to the intermediate school. I visited Avon Falls seven times and was a guest at an all-staff lunch meeting hosted by the science department.

The superintendent of this district was Bob EnSpania (BE), who worked for years as a manager in the public and private sectors before earning a doctorate in educational administration. After working as an assistant superintendent for a large district in the eastern part of the state, he became superintendent at Avon Falls in 2003. I met with him several times, originally at the home of a university professor. His wife, who was a doctoral student, worked in the same lab that I did for a period. In that first meeting he compared his experience as a manager with that of a superintendent by saying education is like a business where "people are the product." Bob facilitated the participation of Avon Falls and provided useful district-level perspectives on assessment in the schools and with the state. In one of our meetings he elaborated on this perspective and related it to the state test, as Transcript 7.4 shows.

Transcript 7.4 - Bob Enspania discusses goals of education and the state assessments.

1)	BE: I never took a psychology class, I never took philosophy class until I had to take this philosophy of education. And I did this, I can't remember who I did this paper on but I did a parallel between education and business and the uh the product for education, in my view the product, is nurturing the inquisitiveness in .. kids to young adults so they continue to want to learn. And I think we don.. you don't have to agree with my definition.
2)	PP: No it is interesting.
3)	BE: But if that is the definition then we do we do a lousy job. My kid graduated from college. Ah why don't you explore this? *Nope I want to be done. I got x number of* Ask everybody..*I want to be done y'know* *I got this masters degree it's 24 credits.* What about this? this is pretty interesting... *nope it doesn't fall within my 24 credits.* Everybody wants to be done. And my view is you need to be able to..you're never done. But uh but when you when you talk about the state..What do those scores mean anyway?

Janey Fess (JF) was the only curriculum director at Avon Falls and is responsible for all subjects and grades. In contrast to the districts surrounding Corrine Eaton's MS Center, Avon Falls had a specialist in Janey who knew science and had taught. She began her career over twenty years earlier as a home economics teacher before becoming a science teacher and later a middle school principal. She left a position as a science consultant at the local ISD for the job with Avon Falls. Although initially appearing cautious about the study, she became increasingly comfortable and engaged in the process.

Janey was conversant in educational theory and her office contained many contemporary research books about various subjects, including literacy and math. In her conversations with me, she often used concepts common in research when discussing students, for example, using the term *misconceptions* as she did in Transcript 7.4. She also used the terms *prior beliefs* in the process of *sense making,* terms that are also common in research about students, to frame her observations about the adults in the schools. She said she was attempting to develop a consistent educational program across the schools and subject areas. She frequently included other subjects in her discussions of science, as illustrated in Transcript 7.5. She also discussed her observations of building-specific thinking, in which the culture of each building is exhibited in the opinions of its staff.

Transcript 7.5 - Janey Fess discusses using assessments to have conversations around instruction.

What I envision is that ... um whatever the assessment is ... that we start to bring teachers together and we decide we want to take a look at an essay that kids wrote in science.

And the two teachers get together at the middle school uh and they gave the same assessment and the same essay question.

And they've decided ahead of time .. they've agreed upon how they're going to score it that's another thing so they know how they're going to score it.

Now they come together and they start looking at what does this tell us about what a kid knows... not that we give them four points or five points but in this I can tell that a child knows photosynthesis but just couldn't get the words out they didn't have the language to do that or another child used all the right y'know .. buzzwords but **really** doesn't understand that plants make their own food when they do that.

So that teachers can start to really look at misconceptions .. and address those and then when they're teaching see they'll start to say, *Oh, I need to be sure to do to deal with this.*

Unless you spend time having conversations around assessments you don't know what to do next .. to inform your teaching.

The principal at Avon Falls is Stan Dubovsky (SD). He is 53 years old, with a slim build and soft manner. He had been a math teacher and then assistant principal before coming to Avon Falls 12 years earlier. His interactions with the staff and his colleagues at the administrative office showed an alertness and quiet engagement in the conversations. Pleasant and laconic, his appearance was informal. I observed him on different occasions in jeans and business casual attire that included wearing a shirt with the school logo to a meeting with me. I observed him in his office suite and standing in the middle of the hallway greeting students walking past at the end of the school day. He seemed concerned with the emotional health of his staff. One of the most animated times in our conversations was when he told me a story about some things he did to boost staff morale. This episode, which I call the “manly men project,” will be discussed later in Chapter 8. In Transcript 7.6, he discussed his perspective on accountability.

Transcript 7.6 - Stan Dubovsky discusses NCLB and MEAP.	
1)	PP: How is No Child Left Behind affected your job?
2)	SD: (5 sec. pause) I think its created a sense of urgency and you in some cases that sense of urgency was needed and in some cases ... it wasn't Its an awful .. lot of .. hoops to jump through even in checking test scores and making sure everybody's taking tests and making sure. As an example on the MEAP test these days..during that two or three week period of MEAP. Because of our setup we proctor almost the makeup exams...And we really go a tremendous amount of effort to make sure we get our 95% of kids tested where as before we would do a PRETTY GOOD JOB (laughter) but it wasn't like we would call a kid in if he was sick and say "Can you come in for an hour and take this science...finish this science test please." And even things like we're always gonna give the reading test and math test or the English language arts and the math test first Science always complains because they’re the last ones and social studies are the last ones to take the test We're always gonna take the English language arts first so that if the kids are gone or on vacation there's some extra time It almost science feel like second class citizens.

His tenure at Avon Falls had included budget tightening, as the school went from what he called a “deluxe team teaching model” with two planning periods per day to a leaner approach with only one planning period. He has also been trying to institute one set of lesson plans for the two teachers in the school and believed that they were not much aligned in terms of topics. The analysis in Chapter 5, specifically illustrated in Figure 5.8, confirms his suspicion.

One other teacher at Avon Falls who was not a case participant was Jonathan Brunson (JB). Jonathan left a career as a manager in an automobile-related industry to become a science teacher. He had been teaching eighth grade earth and physical science for several years and participated in the science department's review of their MEAP results which I discuss in the next chapter. He also reported that he purchased a scanning machine and test sheets for his own use in the classroom. In this way, he was like Faith Churchill and Christy Connolly in integrating classroom assessment technology into his practice.

The Case of Crimson Middle School

Crimson Middle School (Crimson), the school where Christy Connolly (CC) and Ben Raminskis (BR) taught, is located in an affluent, technologically sophisticated district. Crimson had a one-to-one laptop program, and Christy and many other teachers used the Moodle online course support system.[31] However, within its district, Crimson was unique in being the only middle school to qualify for Title I funds based on its poverty level. Furthermore, all of its feeder elementary schools also qualified for Title I, and they were the only elementary schools in the district that did so. At Crimson, school principal Ed Bedminster (EB) reported that 65% of the student population was non-white; 35% of students were receiving free or reduced lunch; 48% of the student population turned over annually; 14% of students were receiving special education services, and 9% of students were English language learners. The school population was 30% African American.

Ed Bedminster is a tall, slender African American man in his early forties. Whenever I saw him, Ed was a man in motion. He was often with kids and exuded a strong, positive presence of leadership in the school. Outside his office door was a five-foot tall portrait of his head with his name and the appellation "Visionary" under it. He reported that his interest in education began when he participated in a mentoring program for black youths in Baltimore. The program (long since discontinued) had black men serve as role models teaching black male students. After this experience, he spent seven years in the classroom teaching both fifth grade and middle school English and math. He

[31] Christy indicated that in the second year of her work with Moodle, usage had diminished from the initial reports to only a core group of teachers. The district also experienced some issues in the availability of computers in the second year of the one-to-one laptop program.

had held the position of either principal or assistant principal for seven years when we met. He reported that he joined Crimson to help facilitate a turnaround of the school, which had poor morale and low test scores. In Transcript 7.6 he discussed his attention to kids' science projects and supporting teachers.

Transcript 7.7 - Ed Bedminster discusses kids working on projects.

1) PP: So when I was here last week I noticed that Mr. Broom was coming in to get names from Mrs. Connolly about students that hadn't completed the..
2) EB: Projects..that's what we do here. If you don't completed the project you're staying here with me and him until you get it done.
3) PP: uhhuh
4) EB: Period. It's the expectation that every kid does it.
Now if they skip it then we give them an in-school you get it done within school. *You've had three months to work on it.* It is the expectation that they do it and it is the way to support our teachers.

Regarding NCLB, Ed was a strong advocate. He said he felt schools had historically ignored black youth, but NCLB was forcing schools to pay attention. I chose not to develop a case for Ed because I had a limited time with him and a small amount of data. While he said he wanted additional data of any kind to be able to help manage teachers, the nature of the school system allowed most teachers to ignore or pay little attention to the MEAP and NCLB. Ed told me that the union system that he operated under gave him very little leverage over tenured teachers such as Christy Connolly and Ben Raminskis. He was interested in assessment and developed a requirement that teachers put copies of their tests and quizzes in a common file that he could access, but the teachers told me it was of marginal relevance. Teachers at Crimson provided the fewest examples of assessments of any school in the study. Ed also indicated that the district provided him with information that compared grades and MEAP results, but that at the district level there were not any formal or informal processes in place to help principals across schools work with the information they had.

The Case of Dorchester Middle School

Dorchester Middle School (Dorchester), where Faith Churchill (FC) is in her 30th year as a teacher, is in a sleepy and affluent rural community in the corner of the large county where Crimson's district resides at the other end. These two schools were the only ones in the study served by the same ISD. The town of Dorchester, much like its name, exudes a sense of prosperous agriculture and includes a picturesque village center with a river running under the Main Street. The Dorchester Community School District at the time of this study was only slightly larger than Swallow Public Schools', with about 1300 students in three school buildings. Almost all of Dorchester's students were white, with about 6% percent Hispanic, and 15% received free or reduced lunches. Based on reported figures for migrant workers, many of the lower socioeconomic students probably have parents connected to farm labor.

The four-grade (5-8) middle school was neat and clean. When I conducted the study in early 2007, principal Donna Dinard (DD) was in her first year in that role. She had come from the high school, where she was a tenth grade English Language Arts teacher. Because of her recent appointment, and also because I was unable to observe any processes related to the use of MEAP or other assessment information, I decided not to develop a case for her. She did, however, express support for the use of information generally in her role as an instructional leader and even introduced a test scanning machine to the school that several teachers, including Faith Churchill, were using. The GradeMaster scanner Donna brought into the school was purchased by her husband (serving in 2008 in the Middle East with the National Guard) for her to use in her classroom. As a literacy teacher, she said she found it was very helpful in automating routine and lower-level tasks. The school had recently begun to use an assessment system from the Northwest Educational Association (NWEA) under a district contract, but only for math and literacy.[32] Donna reported that multiple information sources, including teachers' grades, MEAP, and NWEA, were helpful in decision-making processes where data played a role. However, Dorchester did not use any assessment beyond the MEAP for science.

<u>The Cases of Hardy Middle School, Bob Senoff, and Diane Vander Miller</u>

Hardy Middle School (Hardy) is in a small industrial town that is a major port on one of the great lakes. The town had benefited from both shipping and the distribution of

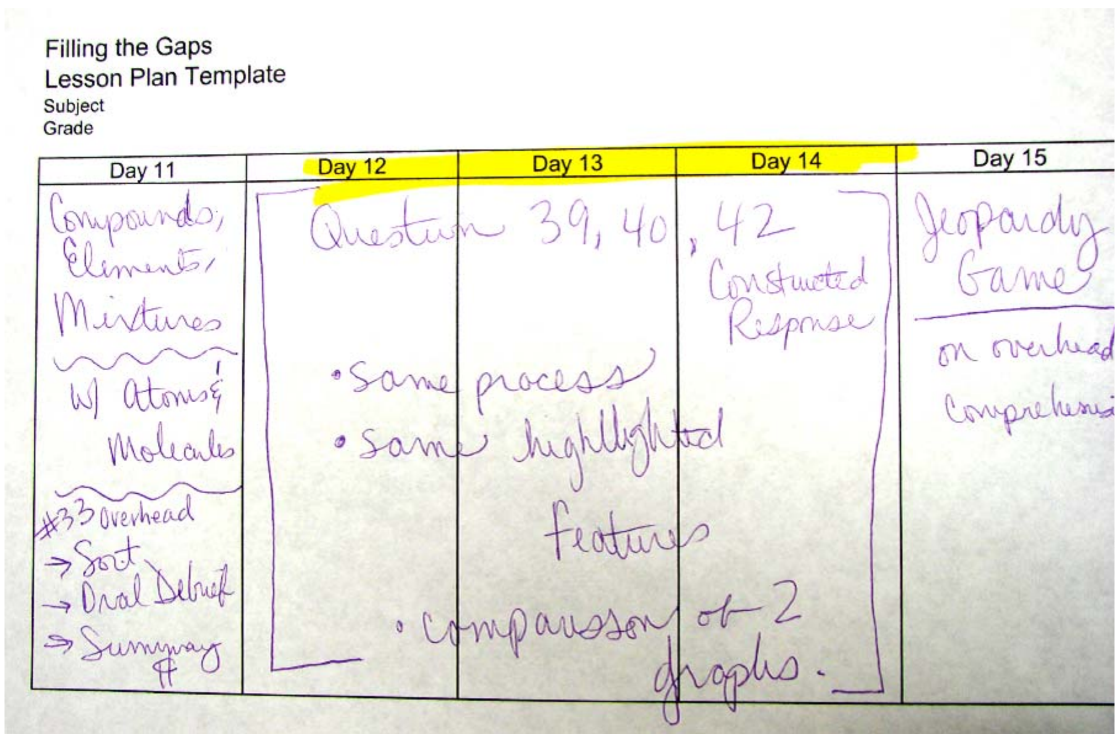
[32] At the time of this study, NWEA had only recently begun to offer science assessment products.

automobile-related manufacturing throughout the state, but at the time of this study was suffering under the same poor economic conditions as much of the rest of the state. The per-capita income in the district was just over $20,000 annually. There were almost 10,000 students in the district, with about 90% of them white. Just over ten percent of the households were below the poverty line.

Hardy was reported by the participants to have been one of the five largest middle schools in the state. In a building the size of many high schools, Hardy was the main middle school in a district that had three high schools, seven elementary schools, and just two middle schools. With more than 1,000 students, the school was organized for sixth and seventh grade into teaching teams, a structure that could have been used to ability group (track) students. Informal comments by some participants indicated that tracking might have occurred. The eighth grade was operating from a curriculum perspective as a school within a school. A Hardy participant reported that the eighth grade is supposed to preview high school for the students, so even though it was within the middle school, it was conceptually different. However, as the discussion about Valerie Jones and the two eighth grade science teachers in Chapter 5 showed, the science teachers did collaborate across seventh and eighth grades despite a perception in the district that the grades were different.

There are two individual cases associated with Hardy. The first case is for Bob Senoff (BS), PhD, who was the district's director of assessment, evaluation, and federal/state programs. He was late middle-aged and of quiet demeanor. Bob was a former teacher, basketball coach, and school principal. I met him at a state MEAP review meeting, where he was serving on a committee reviewing for bias and sensitivity. This committee is discussed in Chapter 8. I then subsequently met him at the annual state student test conference, where he was an attendant. He was instrumental in bringing the MI-Tracker data warehouse system into the district the year I studied. Bob also served on the OEAA technical advisory board.

The other case is for Dianne Vander Miller. Diane was a literacy specialist the year before my study. In the year after I observed her (2007), she was promoted to assistant principal. Her role when I observed her and studied her work was a blend of these two positions. She was brought in to facilitate a major cultural transformation of the school and help teachers to focus on students' performance, with specific attention to the use of data and standards. Bob Senoff had been instrumental in her coming into that position and had encouraged her to participate in my study. Since the principal, Carmen DuRonder (CD), had been in the school for only five years and faced a challenging personnel situation, one of Dianne's responsibilities was to help Carmen and her two assistant principals to become instructional leaders. Carmen was an informant for this study, but Dianne served more as a leadership level case participant. Later in this chapter I use Dianne to represent the school principal perspective for Hardy, as her role was at that time largely leadership and her views were echoed by Carmen in my brief interview with her.



Filling the Gaps
Lesson Plan Template
Subject
Grade

Day 11	Day 12	Day 13	Day 14	Day 15
Compounds, Elements, Mixtures w/ Atoms & Molecules #33 Overhead → Sort → Oral Debrief → Summary	Question 39, 40, 42 Constructed Response • Same process • Same highlighted features • Comparison of 2 graphs.			Jeopardy Game on overhead Comprehensi

Document 7.2 - Page of Dianne's planning sheet for science showing MEAP questions for review

Dianne developed sets of notebooks for the principals with detailed information about what each teacher was supposed to be teaching and when so the leaders could communicate more knowledgably with them. The math and literacy portions of these leader notebooks were detailed to the point of specific objectives and assignments, while the science portions consisted only of the science kit rotation schedule. For the teachers, Dianne developed supplemental materials to assist them in their teaching. She organized the rescheduling of an extra class period that had been used for watching movies and doing homework to be used for standards-based instruction. Dianne also scheduled for the teachers a series of workshops to improve instruction and led teacher teams in collaborative planning (see Document 7.2). Much of this planning was data-driven and tied to MEAP results. As did Janey Fess at Avon Falls, Dianne spoke of the relationship between literacy and science, and she emphasized the importance of students doing well on constructed response items. In Transcript 7.7 she describes her perception of this relationship.

Transcript 7.8 - Dianne Vander Miller discusses relationship between literacy, science achievement.

1) PP: Have you been able to look at that from an informational perspective?
Um. have do you have any place where you have brought the data together and you can say look here this is the relationship?
2) DV: There? I have not done that formally. What I did do at a staff meeting was um pull random children who were a 3 or a 4 on the science portion of the MEAP and have people then just yell out their names and then pull up their ELA score and show that the correlation was very similar.
3) PP: Interesting
4) DV: We do not we do not have a very high number of children for whom they pass the science MEAP but do not pass the language arts.
5) PP: Really?
6) DV: Yeah. And I just haven't had time to do more of a correlation study because because there are a massive number of things going on in this building right now.
7) PP: Sure
8) DV: But it can be the opposite where a child scores .. relatively well in language arts but not in science.
But rarely is it that a child is great in science but not a reader.
Because even our state assessment is mostly .. reading for science. You know, you read it you analyze it you answer the question.

There were a number of other informants in the Hardy study, including Ann Pbzerniac (AP), a math teacher who attended one of several workshops on the MI-Tracker software package that I observed. Ann was a close associate of Dianne’s and formerly taught both math and science at Hardy. Also included as an informant, but not as the basis of a case study, was Mike Wolson (MW), a representative (and co-owner with his wife) of the company that sold the MI-Tracker software. Mike managed the conversion of the state MEAP results into the MI-Tracker package and led workshops with groups of teachers from different schools in Hardy's district. Another important informant in this case was Debbie Huston (DH), who operated the small science center and kit operation in the Hardy district.

The Case of Nancy Newman, Returning Focus to the Intermediate Structures

The two cases of Pete Darmond and Corrine Eaton discussed in Chapter 6 showed individuals whose positions provide them with vantage points to view many schools in many districts. This next case is for Nancy Newman (NN), the Director of Curriculum at the Wushutunesa Intermediate School District (WISD). In this role, she, too, has specific knowledge about district level activities related to the 90 school buildings in ten districts serving 48,000 students.

Nancy is middle-aged and soft-spoken, but not shy. I observed her ask questions in both state forums, like the state testing conference, and national venues, such as the American Educational Research Association conference. With an undergraduate degree in botany, Nancy began her career in education as a secretary, technology trainer, and media specialist in the ISD before returning to school for an educational Master's degree. She then worked in a private school for several years before rejoining the ISD and moving through positions of responsibility until she became responsible for curriculum and instruction. She began pursuing a PhD as part of her preparation for future career steps and was a classmate of mine.

In our conversations, Nancy said that Spillane's (2006) work on distributed leadership had helped shape the way she structured a relationship between the schools and the ISD. Under the approach she developed, teams from a school would apply to the ISD for funding and then commit to work collaboratively to address some problem, such as adolescent literacy. Consistent with Spillane's research, this program was open to informal as well as formal school leaders. Many of the project arrangements that school teams created with the ISD under this program involved data (alas, none with science). Nancy's approach to developing information literacy, she said, was initially informed by the DataWise work of Boudett, City, and Murnane (2005).

Nancy illustrates some of the types of institutionally structured network connections that exist in my study. She was a member of the MSTA and was still on its listserve in 2007. She was a regular attendee in the Michigan School Testing Conference and was scheduled to present there in 2008. She had also participated in and presented to the annual administration meeting of the Michigan Association of Intermediate School District Administrators under the sponsorship of her boss, who began serving as president of that association in the 2007-2008 school year.

Since Wushutunesa is the ISD for both the district that Crimson is in and Dorchester's district as well, Nancy was also an informant on those cases. She provided valuable complementary perspectives to those of the individuals already discussed. For example, Nancy told me that she was responsible for the technology program at Crimson that gave each student and teacher a laptop. She was also the management mentor for Donna Dinard, the new principal at Dorchester Middle School. And she provided important insights on what happened to the MS Center that Pete Darmond mentioned earlier in Transcript 6.2 as having suddenly disappeared from the program. As we can see in Transcript 7.7, she was involved in that decision, and she describes some of the motivations for the change.

Transcript 7.9 - Nancy Newman discusses the Math/Science Center (from two portions of interview).

1) NN: You know we have our Math Science Center so we've abdicated a lot of our responsibility for math and science to our Math Science Center which has traditionally meant that is has been abdicated to this one particular .. whoever is the director of the Math Science Center and so uh that is a responsibility we're um taking. I am taking [it] back now.

---

2) NN: I actually have a philosophical problem with the whole fact that we delegate it to the science center.

3) PP: you do?

4) NN: Yeah, and that's part of why...what's happened last week is that our ISD is taking that science center and we're going to run it out of here because we have a very team-based approach and it just drives me insane when we have silos. And it's a silo to me.

In these short excerpts from our conversation, Nancy indicated that the culture of the science center was distinct from the "team-based" approach that the ISD wanted (turn 4). The term "silo" is one that has come to mean an independent management/information structure that is inaccessible by the rest of the organization. It was the subject of a popular book by Harvard professor Patrick Lencioni (2006) in his "Silos, Politics and Turf Wars: A Leadership Fable About Destroying the Barriers That

Turn Colleagues into Competitors." She indicates that rather than abdicate this responsibility, she was taking it back into her curriculum area (turn 1).

Sketching the Network

Viewing the cases presented thus far - the teachers presented in Chapter 5, the science teaching community in Chapter 6, and the MEAP development process, along with the additional cases for schools and individuals given in this chapter – it is possible to represent them as a web, with a chain of relationship linkages from classrooms to the state government. What was initially described in this study as a vertical sampling has been shown through an actor-network analysis (Latour, 2005) to be a set of related entities that I represent in Figure 7.1. This illustration allows for the roles of boundary spanning individuals (Tushman and Scanlan, 1981) to be seen. It does not imply that these paths are balanced or evenly distributed across the cases. Rather, this figure shows that in the data presented, different types of relationships with differing strengths and influences are identified.

The network topology in Figure 7.1 shows how during this study, some individuals in some institutions had greater or lesser access to various resources for science education and assessment than others. In addition to individual connections, it is also possible to identify the role that technologies and materials play in linking organizations.[33] In this conceptual schematic, the lines connecting individuals and organizations vary to indicate strength (thickness), currency or how active the relationship is in the present (current is solid and inactive is dotted), and direction of influence (arrows). Some individuals, such as Pete Darmond, are highly involved with the community and influence its operations. Others, such as Faith Churchill, tend to be more recipients of changes in the community.

[33] Individuals also have roles in the technological and material relationships, but that level of detail is not included in this study.

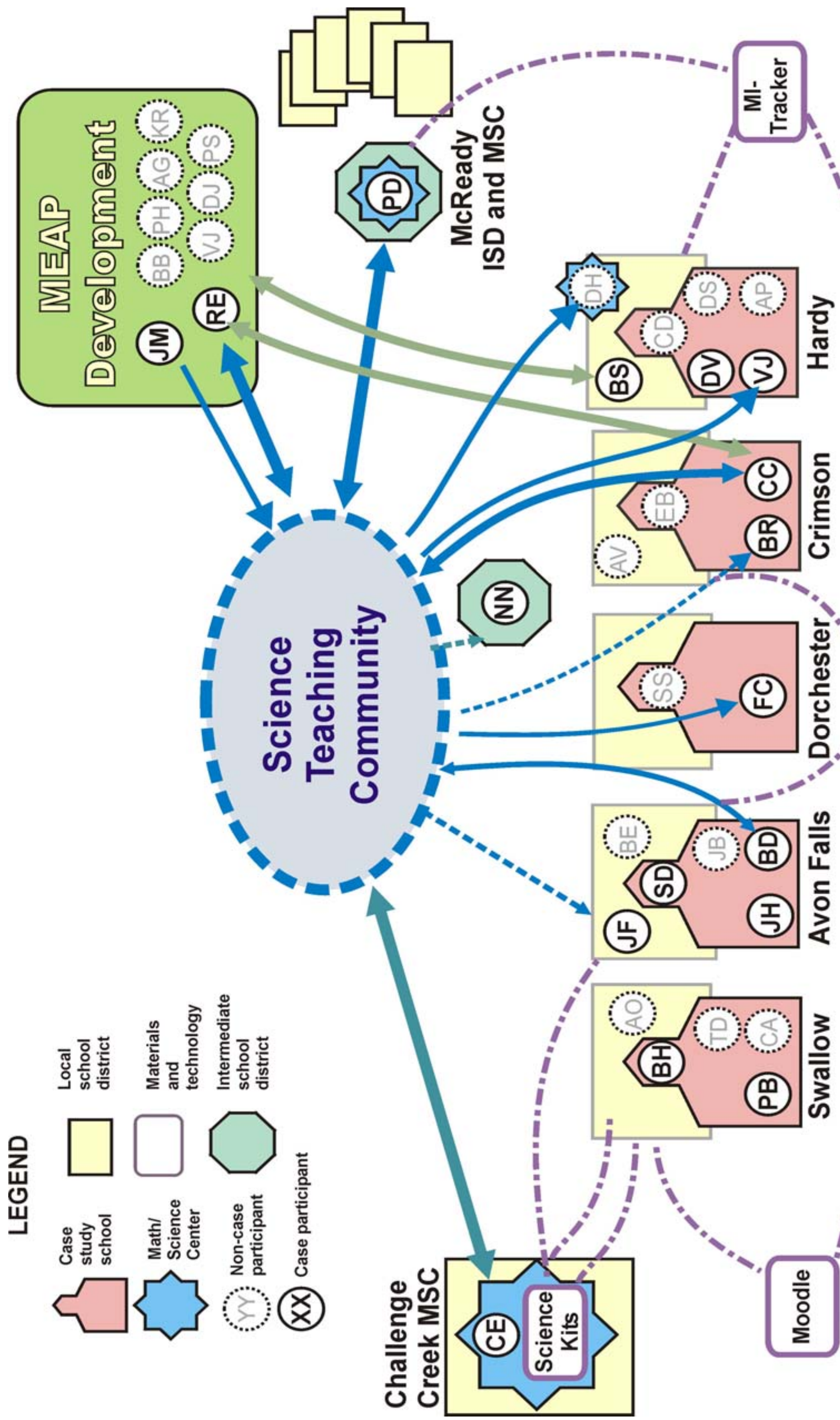


Figure 7.1 - Network of relationships related to science education and assessment

Some individuals appear as brokers (Wenger, 2000) through institutionally structured relationships, while others do not seem to fit that classification. In Swallow, for example, neither the teacher nor the principal were members of the MSTA and neither had served on state test development committees.[34] Conversely, at Crimson, Christy Connolly was active in MSTA and served on the state committees convened to review the state test. The individual's role in these institutionally structured relationships does not seem constrained or influenced by the school, since both Jim Heinrich and Ben Raminskis were not involved in MSTA while the other teachers in their school were. Furthermore, these institutionally structured relationships can be seen as either unidirectional, as is the cases of Faith Churchill or Valerie Jones, or dialectic, as the cases of Pete Darmond, Corrine Eaton and Christy Connolly show. In the next section of this chapter, I use Table 7.1 (appearing several pages below) to summarize these brokering relationships, providing a rationale for the nature of the individual association lines used in Figure 7.1 and an expanded version in Figure 7.2.

Figure 7.1 also shows three different types of material relationships. One involves classroom materials, including textbooks and assessment resources in the examples of the science kits. This category could also include commercial publishers' products, as several of the teachers in this study used the same textbooks, although they relied upon them in very different ways. While this aspect was not studied, these materials are often associated with professional development activities that may bring individuals together and foster common practices. The second type of material relationship is classroom technology, in the example of the Moodle online course management software. The third type is represented by the Successline MI-Tracker data system, which was largely a back-office technology designed to manage information that primarily comes from assessments, organized by state standards.[35] Not illustrated, but conceptually part of the category of classroom technology, are the automated classroom data collection technologies used by Faith Churchill in Dorchester and Jonathan Brunson in Avon Falls.

[34] Other science teachers at Swallow were reported to have been MSTA members.

[35] The software vendor develops a custom version for each of the several states in which it it operates, so that the reports and functions in the software are represented according to state standard structure.

Two More Professional Associations and Representative Participants

While the science education community is the most important community for this study, I also analyzed two other professional associations that make important evidentiary contributions. I included them in the study, in part, to pursue the question of assessment systems that are alternatives to classroom praxis and the MEAP. The discussion of survey responses in Chapter 6 left open the possibility that such a system could have been in use for middle school science somewhere in the state, even though no direct evidence of such a system appeared. These two professional associations help explore this issue, and later in Chapter 9 I develop an explanation for the apparent dominance of the two types of systems managed by teachers and the state government.

The School Administration Community Case

The first professional association in this study for science teachers was presented in Chapter 6 for the. The second professional association is the association for school administration. This category included two organizations that shared the same offices and some staff members: the Michigan Association of Intermediate School Administrators (MAISA) and the Michigan Association of School Administrators (MASA). These organizations are among several that were formed to advocate for and serve those responsible for school administration.

Document 7.3 - MAISA Web Home Page

Other organizations that could be found in the same building or nearby were the association for school boards and the association for Michigan School Business Officials (MSBO). Also in the same office space was a non-profit organization created by MASA, MAISA, and MBSO called the Michigan Institute for Education Management (MIEM), which was one of the sponsors of the annual MSTC testing conference that will be discussed in the next case. Similar to the way that the MSTA became the signature organization indexing a set of affiliate groups, MAISA will be used for convenience in this dissertation to refer to the professional organizations that serve district and intermediate district school administrators.

Document 7.3 shows the MAISA home page, with certain aspects regarding this study identified.[36] The relationship to MASA is indicated with a direct home page link (a). MAISA's role in disseminating information about the MEAP and state standards is shown by a lead story (b). Three indicators of change occurring in the state that will be germane to the discussion in Chapter 9 are shown in the data warehouse surveys (c), the

[36] I modified this image from its original to better fit the page. Originally, what shows in this figure as two columns was in fact one longer column.

discussion of the funding crisis (d), and the information on Public Act 63 (e), which directs ISDs to look for and report on opportunities to collaborate and consolidate services. The documentary evidence for this case is distributed over many types of texts. MAISA does not publish a newsletter or journal as the MSTA does. But it does publish minutes from a variety of committees, as well as position papers. MASA publishes a bimonthly newsletter that I reviewed from 2003 through 2007.

Transcript 7.10 - Karen Minor discusses how the DATA4SS project came together.

1) PP: In your experience, have you seen projects like this at the state or is DATA4SS kind of different for some reason?
2) KM: I believe it is a little but different project most that have occurred and I think part of that is timing and part of that is players ..um ..the executive committee a number of us serve at the state level.
3) PP: OK.
4) KM: Frequently, I serve on advisory team for OSI I've served on work groups and advisory groups for OEAA.
5) PP: OK.
6) KM: So know both those groups, I have personal relationships with both those groups.
   We have Betsy Rondo who is a key legislative connection. She works with a number of organization. Cathy Swallow in McReady has been influential in highly involved in rewriting some of the data sets for SRSD and CEPI's Work. Don Pulte is a leader in MAISA and the state technology committee.
7) PP: Umhm
8) KM: So we have a number and that just every one of the members of the executive committee are involved at the state level as well as at the local level.
9) PP: OK
10) KM: And I think has helped a lot. And I think the state department has done some significant changes in the past .. probably three to five years
11) PP: umhm
12) KM: in how they are approaching working with locals. The uh Mike Flanagan darkening the dotted lines concept, ISDs working in partnership with the state department.
13) PP: Umhm
14) KN: And that's has been underway now for three to four years. So I think they're seen more as partners now than hands off, we're working back and forth much more.
15) PP: uhuh
16) KM: I think that's been a little bit unique. I also think that the fact that we have so many different connections in different departments the state level through the executive committee has been real beneficial. So for example the DATA4SS site some of that the the sample reports were handed out by OSI at the rollout for the comprehensive needs assessment for the title one school.
17) PP: OK
18) KM: And so what we're finding is more and more connections occurring.

The case for school administrator has several informants that reveal not only other aspects of the networks that operate between the classroom and the state, but also more detail on the multiple connections between the individuals who operate as brokers between these layers. Two case informants are Don Pulte (DP) and Karen Minor (KM), who both worked for the Tichiochi ISD. Don was an assistant superintendent. He had been an elementary school teacher with a specialization in science and chaired a technology committee for MAISA that supported a survey of ISDs on their data warehousing efforts. Karen was a curriculum specialist, whose position was roughly analogous to Nancy Newman's. Karen was responsible for a project that began in 2006 called the Data for Student Success or DATA4SS (pronounced *data force*). DATA4SS was a collaboration that included McReady ISD, Cummings ISD (the ISD for Avon Falls), and several other ISDs as beneficiaries, but not leaders, of the project. Karen described its origination in Transcript 7.10.

Karen's discussion of the uniqueness of this project is important for several reasons.   In Transcript 7.9 she provided a rich account of the interconnections necessary for a project like this to occur (turns 2-17) and identified connections to McReady ISD (where Pete Darmond works) and to Don Pulte. She also described how the future DATA4SS information architecture will allow data that schools send to the state to be then sent back to the schools for use, even when students move from school to school or from district to district. This architecture has implications for some of the material connections illustrated in Figure 7.1.

<u>The Cases for Michigan Testing Community and Bernie Lauer</u>

The third professional case I construct is for those participating in activities related to student assessment or testing in Michigan. This is mainly represented by the Michigan Educational Research Association (MERA), which was established in 1972, and the Michigan School Testing Conference (MSTC), which began in 1960. As with the other cases for professional associations, this case draws on other affiliated organizations, including connections to the school administration case. In the expanded network diagram (Figure 7.2) I show these two communities as specifically related.

Bernie Lauer (BL) had been active in both MERA and MSTC for years and played a role similar in some ways to roles that Corrine Eaton and Pete Darmond played in the science education community. Bernie had never been a teacher or “had a real job,” as he said to me. He joined the ISD at which he worked, Maple Public Schools ISD, after getting his doctorate in educational measurement from a state school. He had been president of MERA and a regular presenter for the previous ten years at both the MERA meetings and the MSTC, where I met him after watching his presentation and later joined him for lunch. The MSTC conference programs since 2003, which was the farthest back that I could review, showed Bernie had been a speaker in every conference and a member of the planning committee three times. He was usually involved in more than one session and frequently chaired sessions.

Bernie was an enthusiastic study participant and promoted the idea of using data to support instruction. His long tenure as a leader in this community provided him with the ability to comment not only on what had been occurring within his own ISD, which was responsible for 230,000 students across all grades and 62 middle schools,[37] but also on what had been occurring more broadly in his field. When asked about middle school science assessment alternatives to the MEAP, Bernie indicated that while the possibility existed in the near future, at that time such systems did not, to his knowledge, exist for middle school science.

Over the course of our conversations, Bernie showed his belief in the potential for good assessment information to be used directly in instruction. In Transcript 7.11 he began a talk on his ISD’s technology efforts in a 2007 MSTC session by drawing a distinction between data warehouse systems and student assessment or student performance systems. Later, in Chapter 9, I will touch on the differences between the two and what an appreciation of the differences might mean for what had been occurring within the state.

[37] From ISD website: http://www.oakland.k12.mi.us/AboutUs/WhatWeDo/tabid/138/Default.aspx

Transcript 7.11 - Bernie Lauer discusses his ISD's data analysis system.

I'm Bernie Lauer from Maple Public Schools and we don't think we have a data warehouse.
I mean literally, we do not call what we are doing a data warehouse. You can't put in which bus, well you could, but we're not particularly interested in analyzing which buses a kid is on.

We are not particularly interested in salary schedules.

What we're interested in is student achievement.

And um that's what we focused on.

Now who knows what we'll .. end up thinking somewhere down the road but I'm just gonna describe to you where we are and what we've done and what we've learned.

We have called what we have done a data analysis system and a student assessment system.

The testing community case introduces not only new participants to the study, but also new connections to existing participants and another lens on the network properties of educational organizations. Nancy Newman, as a curriculum director, had been both a regular attendee of the MSTC and a presenter. Bob Sendoff was a member of MERA, and both Bob and Bernie were members of committees that advised OEAA on various matters related to the MEAP program.

When these new communities are added to the network diagram, a richer illustration of the types of links between various types of actors emerges. This more complete representation of the study's evidentiary web shows that at the middle levels, these irregular organizations between the schools and the state are situated some highly networked individuals with multiple institutionally structured connections. These connections allow them to participate not only in hierarchical activities, for example as an ISD staff member to support the schools within that ISD's jurisdiction, but also to participate in activities that link middle layer organizations and relate to even more schools. This participation suggests that they can appreciate a broader range of issues in the state than may be afforded to less connected actors.

Table 7.1 summarizes different brokering/boundary crossing activities by individuals who were discussed in both this chapter and the preceding one. Not all individuals are listed, but only those where I have been able to identify activities that they engage in, or in the case of a few, that they previously engaged in related to an organizational case in this study.

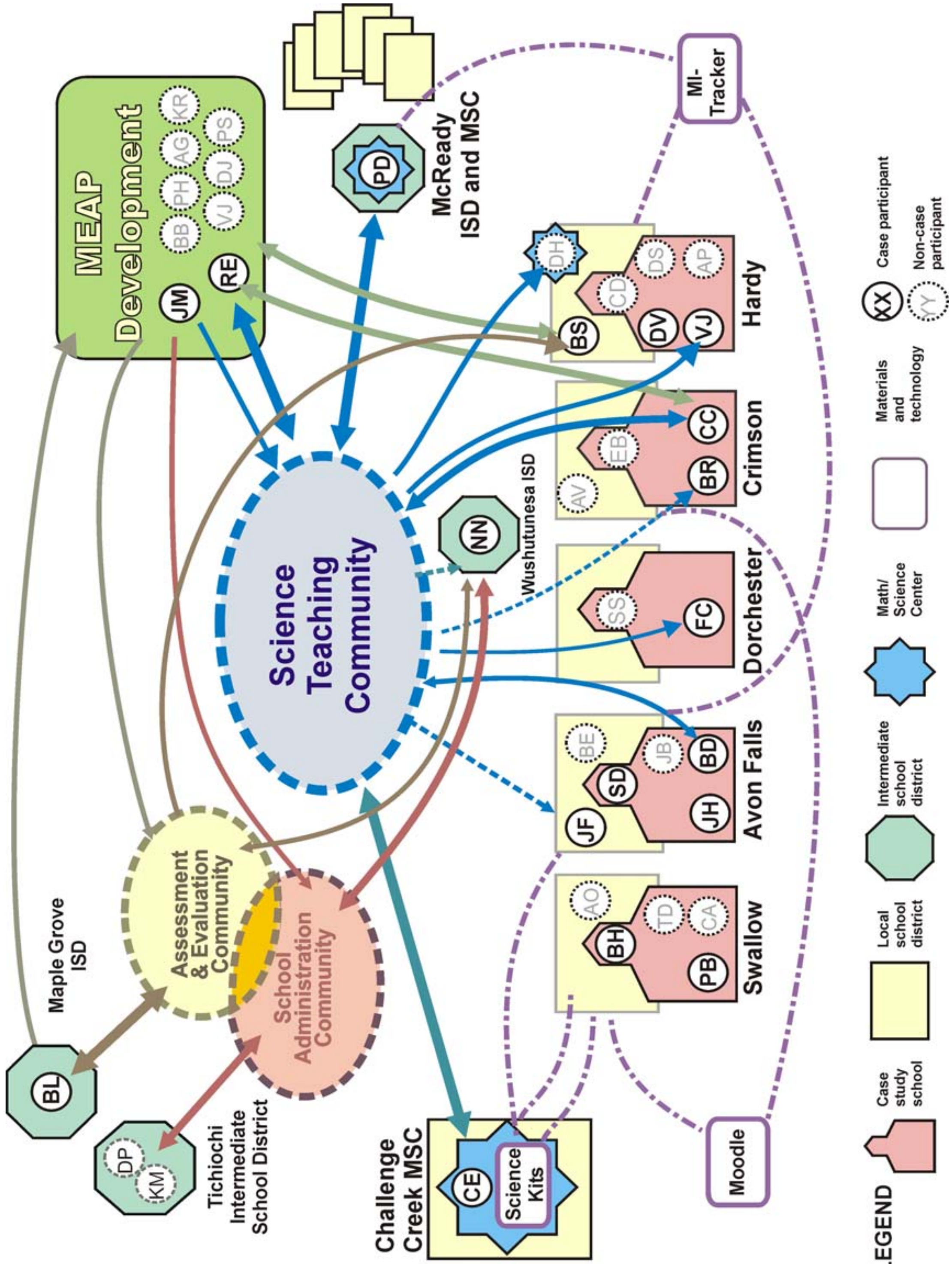


Figure 7.2 - Full network representation for the study.

Table 7.1 - Summary of brokering relationships in the dissertation's evidentiary web

Individual	Case	Identified Boundary Crossing
Joshua Martinque	Assessments	Frequent presentations
	Science	Occasional presentations
Roscoe Ellis	Assessments	Regular presentations
	Science	Regular presentations, writing for journal, recruiting science experts
Nancy Newman	Administration	Presentations
	Assessments	Presentations
	Science	Former member
Bernie Lauer	Assessments	Leader, presentations
Pete Darmond	Science	Leader, presentations, writing
Corrine Eaton	Science	Leader
	MEAP	Committee member
Bob Senoff	Administration	Member
	Assessments	Member
	MEAP	Member of review and steering committees
Janey Fess	Science	Former member
Christy Connolly	Science	Member, presenter
	MEAP	Committee member
Betsy Dearing	Science	Member, presenter
Faith Churchill	Science	Member
Valerie Jones	Science	Member
Ben Raminskis	Science	Former member

Cross-Case Analysis of Beliefs Regarding Assessment/Accountability

Having illustrated the relationships between these cases, I began to look across them for possible patterns. Some of the research discussed in Chapter 2 looked at multiple levels of the educational system, including Coburn and Talbert's (2006) study of districts and schools; Massell's (2001) study of several states and their schools; and Ingram, Louis, and Schroeder's (2004) study using a Total Quality Management framework. These studies reported results at different levels of the system that indicated teachers and district administrators have different values for evidence and assessment. To explore how this issue relates to the network of participants in this study, I look at how they displayed opinions regarding high stakes assessment and accountability. The participants' presentations of their positions on accountability are summarized in Table 7.2.

I re-present their positions here with the understanding that in many cases, the individuals were not specifically interviewed about epistemologies regarding assessments. Even if they had been, as Ikemoto and Marsh (2007) and Ingram, Louis, and Schroeder (2004) illustrate, responses to such direct questions may often be unreliable and tend to affirm pro-data, pro-accountability positions that may not be consistent with the respondent's actual practice.

Table 7.2 - Summary of case participants' presentations regarding accountability and assessment

	Case Part-icipant	Positions presented about accountability: Positive	Negative	Comments
MDE	JM	assumed	strong	Concerns about fairness and technical accuracy.
	RE	assumed	--	Advocated good assessment.
Mid-level	CE	--	--	Presented other forms of classroom information.
	PD	--	--	Recognizes importance of the issue.
	BL	strong	--	An advocate for evidenced driven instruction.
	NN	moderate	minor	Presented alternative assessment/information.
District	BS	assumed	--	A member of OEAA committees.
	BD	--	minor	Considered NCLB analogous to OSHA.
	JF	moderate	--	Positives associated with conversation.
Principals	BW	minor	moderate	Active in using/presenting results.
	SD	moderate	moderate	Very balanced views.
	DD	moderate	moderate	Acknowledged as appropriate but burdensome.
	BE	strong	minor	Believes in NCLB principles but not enough data.
	DV	strong	--	Strong believer in accountability and equality.
Teachers	PB	--	strong	Sees accountability as an irritation, not relevant.
	FC	strong	minor	Mostly supports, considers burdensome.
	BR	some	some	Active in trying to use MEAP results.
	JH	--	strong	Very negative. .
	CC	strong	---	Very strongly positive.
	BR	minor	minor	Very balanced views.
	VJ	--	minor	Seemingly not a major issue.

Given the previous research on belief systems discussed above, I included three questions in the structured interviews designed to provide opportunities for the participants to share their opinions on the subject. These questions were not direct and did not ask whether they agreed or disagreed with accountability, because many of them were required by their job to be proponents. By asking questions such as, "How is NCLB affecting your job?[38]" I provided an opportunity for participants to present their views without directly confronting them about their commitment to policies to which they were all expected to conform. Their views were also indicated in response to other questions, such as the one I asked Burt Wainwright in Transcript 7.2 about how assessments could help his job. His response reflected concerns about the current implementation of accountability testing and its implications for good instruction. Regardless of the questions asked, all of the participants' responses were evaluated as a group, so that I

[38] This is also identified as RQ12 or research question 12 from the questionnaire.

could include the broadest amount of evidence rather than only those responses to specific questions.

This comparison of positions across different types of roles indicated that the participants in this study did not support a leveled model of belief systems with regard to something as important as accountability and assessment. While there were teachers (ex: Paul Bond and Jim Heinrich) who expressed a contrary view of high-stakes assessment, there were others (ex: Christy Connolly and Faith Churchill) who were strongly positive. At the teacher level and in levels above, many held balanced views, and many expressed sympathy for the fundamental principles of policies like NCLB while also acknowledging difficulties in its current implementation. An important example is Joshua Martinique. As a representative of the state government who is responsible for implementing accountability systems, his job requires him to advocate for the accountability process. And yet, he has made efforts to highlight technical difficulties with equity consequences in the current implementation of NCLB. These representations were made not only to me, but also in publications and conference presentations. He reported that he has made efforts to have the federal government alter its approach to the testing of non-native speakers of English because of the formidable measurement challenges. He is not alone in the state. Bobby Black, the senior member of the OEAA team, discussed with me at length some of the fundamental challenges that exist in measuring learning over time.

Reassembling the System: The Network as Lens

The approach this chapter has taken in representing educational organizations in Michigan introduces new opportunities and complications. When the system can be constructed as a hierarchy, the possibility exists to shift certain analytically relevant issues to higher levels, in the same way that Latour (2005) said sociologists have often pushed important analytic issues to the black boxes of “society” or “actor” without unpacking these constructs to see that these categories are often difficult to directly observe. His solution was to follow the actors, ande considers actors (actants) to include material artifacts, as well as humans. In following his model, this chapter has built a network of evidence that becomes a way to view a broad set of practices, in the same way that an ecologist might construct a food web diagram to illustrate analytically important aspects of an ecosystem.

But in pursuing this path, in leaving the comfort of the hierarchy, some important questions of representation in this model are emerging. In addition to the important and historically established questions about number of participants, from which localities, and with what characteristics they are drawn, there are new questions about the collection of evidentiary loci included in the study and the path of expansion from some starting loci to a larger collection. In addition to asking how the story might be different if different participants (ex: urban teachers) were included, we can ask, how might the story differ if I began with an intermediate layer -- for example, curriculum directors or Math Science Center directors -- and expanded outwards to the state and the schools? In Chapter 10, I will attempt to examine these issues, as well as my own role within this study's network.

Chapter 8 Boundary Practices

"One's destination is never a place, but a new way of seeing things."
~ Henry Miller

Introduction: Activities Across Two Types of Assessment Systems

Now that this study's network of interrelated individuals and organizations has been described and the dominance of two types of assessment systems (classrooms and annual MEAP) for seventh grade science established, I focus this chapter on activities where these types of systems relate. Following Wenger (1998), I will begin by focusing on boundary practices: social events where two activity systems come in contact. The boundary practices I focus on in this chapter come from two locations: the state test development process and case study schools.

There are two aspects of this chapter's analyses that are important to emphasize. The first is that this chapter examines both production and consumption practices related to the annual MEAP state test. That makes this analysis qualitatively different from those studies reviewed in Chapter 2. Those studies all focused on the usage of information and not the design of the instruments that are used to generate that information. An important part of this account of the boundary is accomplished by describing two types of activities. One is managed by the state and includes teachers and science educators from local settings. The other is managed within two different schools where, similar to other schools in Michigan,[39] school personnel are reviewing their annual test results. In presenting aspects of both the creation and use of the state test, the analyses in this

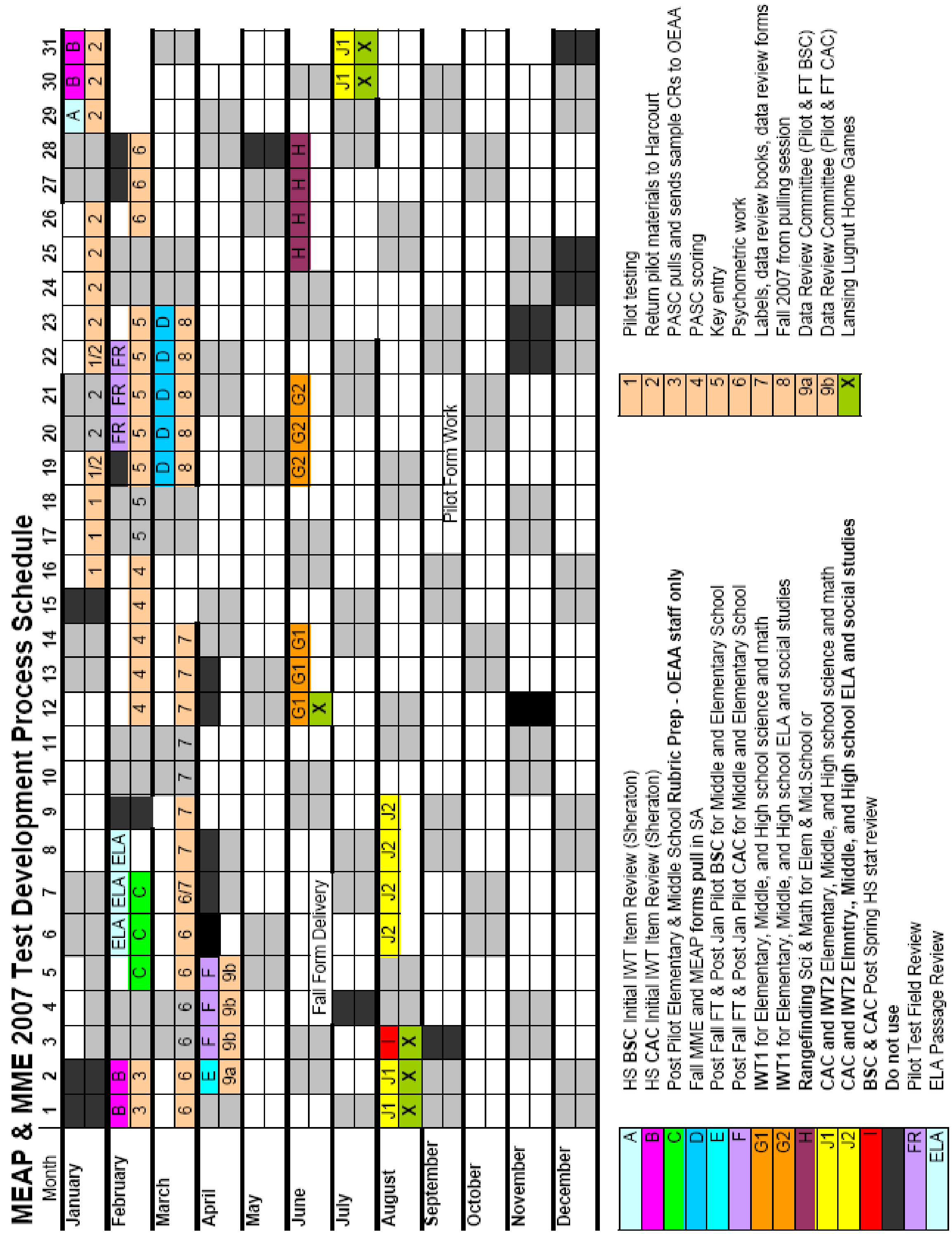
[39] This similarity was established through questions in the MSTA survey discussed in Chapter 6.

chapter develop another type of coherence in addition to the relational linkages identified in Chapter 7. They show a broad structuring across this network that highlights the role of test items as boundary objects (Star and Griesemer, 1989) that complement other theoretical constructs of brokers and boundary practices (Wenger, 1998) in a systemic description.

In this chapter I first discuss the MEAP development process and focus on three specific boundary practice events that occurred in 2007. Next, I review two events that occurred within two of the case study schools, Swallow and Avon Falls. In both of these first two presentations I am sketching a process that occurs over extended timeframes to locate these events within a larger temporal model. After presenting the descriptions of these discussions around MEAP items, I will analyze the topics raised across the five events. This analysis leverages the MEAP item as a structuring entity. Then I shift the focus away from the MEAP Item to three other perspectives of the interaction between the state and schools. The first two focus on schools, using the school cases to contrast leadership responses to the accountability process and then the MSTA survey, combined with school cases, to look at how the MEAP test and results release schedules impact the school's calendar. The final analysis looks at the role of boundary spanners (Tushman and Scanlan, 1981) or brokers (Wenger, 1998) who cross organizational boundaries in the service of science education in Michigan.

Boundary Practices in the State: MEAP Item Reviews

The MEAP development process described in chapter 7 and illustrated in Figure 7.3 is managed by the State of Michigan with the support of a commercial test development vendor. This is a formal process, with steps that are publicly described and subject to a peer review process by other states, as required by the U.S. Department of Education. The process involves many steps, from the initial drafting of test items to several reviews by practitioners. I focus on three types of review processes that involve educators discussing items originally written by Michigan science educators. The goal of this formal development/review process is to develop test items that appropriately measure the depth of knowledge by students and do so fairly across sub-groups.



Document 8.1 - MEAP Development Calendar from MDE.

Document 8.1 shows Michigan's MEAP/MME test development schedule for the time of this study. The process involves several committee reviews. These reviews may occur at two time periods in the life cycle of an item: after the item is written but before it is field tested (activities A/B in Document 8.1) and after the item is field tested (activities E/F/I in Document 8.1). The reviews that precede the field test can be for bias/sensitivity or content. The bias/sensitivity (BSC) committee looks at whether the wording of a question is difficult for a sub-group and whether the visual image(s) in the item can be made accessible to persons with visual impairments. The content advisory committee (CAC) reviews the science involved in the question. The post-field trial reviews are also called "stat" reviews, because the committees are given statistical summaries of student performance on items and the committee's use of that information in deciding whether to recommend changes or to drop the item from the development process.[40] Two other important reviews involve range finding and development of proficiency (cut) scores. Range finding is an evaluation of constructed response question answers to define how raters will assign scores. Cut-scores are used to develop the overall proficiency levels assigned to students based on cumulative test scores.

Photo 8.1 - Review committee (represented in Transcript 8.2) discussing MEAP items.

Note: Participants in composite photo are: R1 (at left). R2-R3. BB , R4 –R6, VN is off camera

Each meeting I observed followed a similar pattern. First, the MDE and test vendor representatives provided instructions to the committee members and gave them packets of materials, including printouts of the items. They also distributed non-

[40] Items dropped from the process were reported to become the property of the test development vendor so that they could be used on other test products. At one point, the vendor representative discussed with the committee an area where his firm had a need for a particular type of assessment item for their inventory.

disclosure forms that all attendees (including me) signed. Because the test items they were reviewing had not yet been used on a test and thus were not released to the public, security restrictions would not allow me to reproduce any of those items or the work documents associated with them. Further, none of the forms that were used by the vendor when working on these items were allowed to be shared, as they represent proprietary procedures of this company.

The reviews then proceeded on an item-by-item basis. For each item there was some discussion about what might be a concern with the item, followed by some decision to modify, reject, or accept the item as it was. The issues raised and decisions taken varied by item and by committee. These item-oriented discussions form the core of my analysis of these sessions because 1) the decisions taken as a result of these discussions directly influenced which items could appear on a test and in what form; and 2) the considerations the committees entertained included a range of systemic issues.

Since the committee members included senior science specialists with both classroom and leadership experience, they were able to discuss the items from a range of perspectives. The topics that were common in these discussions included the nature of the state standards and what the standards were asking for; the ways that teacher could or should be teaching, and classroom equipment. At the state level I recorded and reviewed 64 item discussions across three committees, summarized in Table 8.1.

Table 8.1 - Overview of MEAP item review meetings in study

Activity	Committee Type	Focus	Number of Item Conversations
Bias review	Composite of science and non-science	Reviewing for visual and cultural issues	16 items, mostly graphics and terminology issues
Content review	Mostly science	General design issues	38 items, range of discussions
Stat review	Mostly science	Minor adjustment or drop items	10 items, in-depth discussions

For each one of the review sessions I analyzed, I first segmented the media record of the meeting into discrete discussions. Other than the introductory part, almost all of the meeting was structured around the review of specific items, so each item discussion became a segment of the discourse. For each item discussion I then assigned qualitative codes using the coding scheme described in Chapter 4. As a result, I am able to describe some of the patterns that occur across these discussions later in this chapter.

The first example, in Transcript 8.1, comes from the bias and sensitivity committee. In this discussion, an unnamed committee member[41] (R1) with expertise in issues of visual impairment identifies issues with certain visual choices on the item (turns 2-9). This item has the students select from different representations of organic compounds. This review is early in the development process, so the item can be revised substantially. Roscoe Ellis (RE), who is co-leading this meeting with the vendor representative (VN), asks about labeling of the figures and whether or not the labels can be converted to Braille. Another reviewer (R2) then asks about the text in the image (turn 15). My coding for this item discussion identified it as relating to assessment artifacts (both graphics and terminology). Both of these are distinct codes described in Table 4.5. The decision taken by the committee was to modify this item slightly by making the item image easier to see (turns 6-13).

Transcript 8.1 - Review of item for visual accessibility and conversion to tactile/Braille
1) VN: Thirteen
2) R1: Well the toughest one .. I know these are uh samples but the toughest one to see will be the first one A, it will be the hardest one to tactile graphic.
pause
3) VN: Any suggestions on how to fix that?
pause
4) VN: Yeah because it looks like a three dimensional type of uh .. could you flatten it out?
5) R1: Yea like taking the dimension out would be great.
6) VN: to [distracter] 2b?
7) R1: yea if you could do that for the outside ones. And then looking down at C it just needs its fading needs a black outline the contrasting colors need to distinguish the background. So that's what is going on with that. But you'll do that I mean you already said you
8) VN: I'm sorry, say that again? The black outline?
9) R1: The black outline is like its like kindof squiggly it needs to be defined..
10) VH: Oh I see
11) VN: its fading it needs to be a better contrast.
12) RE: And do the two compounds in D need to be labeled?

41 Anonymous participants who are quoted from meetings signed participant consent forms. I leave them anonymous because their identity is not used in other parts of the study, where named informants often appear more than once.

Transcript 8.1 - Review of item for visual accessibility and conversion to tactile/Braille
13) R1: I was just going to ask that .. yeah because .. especially the glucose at the end that's looking more dimensional the others were OK.
14) VN: So you want us to make the carbon dioxide larger?
15) R2: And make sure that the size of the writing?
16) VN: uniformed throughout yes the labels will be uniformed throughout.
17) RE: Can this be Brailled?
18) R1: Yes I think it can. Yes I think it can be done pretty well. It is just the first one scared me but I think if it will be 2 dimensional it will be better.

The next item review I present involves broader discussions. It comes from a content committee meeting. The key issue in the example shown in Transcript 8.2 is how the question choices should be worded to access students' understanding of the nature of matter, a key middle school science concept. The conversation begins with two reviewers (R1 & R2) discussing whether the question should be worded to ask students to identify elements or atoms (turns 2-5). The terminological distinction is important, as "element" refers to the class of matter, whereas "atoms" refer to the actual physical instances of that matter class. The reviewers wonder whether the term "atom" is allowable under the standards or whether the more general term "particle" is defined as appropriate for this age group. "Particle" can refer to compound molecules (ex: carbon dioxide or glucose) as well as particles that are of one type of element (ex: oxygen). The discussion proceeds to include the state standards, with one R1 claiming that the national and state standards differ on what language should be used. The conversation then brings in another perspective when a member questions how students could get this wrong, as their everyday knowledge would tell them that water is a liquid at room temperature (turn 14). My coding of this item discussion includes that it is relating to student knowledge/cognition in addition to terminology and standards.

Transcript 8.2 - Content committee reviews question 94 on the particle model.
1) VN: OK 94 (these numbers are pages in the booklet of potential items)
2) R2: Can't say element either?
3) R1: Shakes head (indicating no)
4) VN: Good question what's the point of the data table? (reading a comment from another context).
5) R1: Well parts of atoms, no I don't think you can say atom at all
6) R2: Well "best describes"? (reading part of question)
7) R1: No you can say particles at middle school but you can't say atoms. I'm hoping

Transcript 8.2 - Content committee reviews question 94 on the particle model.

that will change.

8) R3: Really?

9) R1: I'm sure when we get to grade level (state standard development process) I'm sure the national standards have it

10) R2: Well nobody picked (choice) A at all

11) R4: That's strange I've never seen

12) R5: That's a problem

13) R2: Well let's see "water, carbon dioxide and sugar"

14) R4: It says room temperature you have it right in the front and obviously all kids are going to know that water at room temperature is not a solid.

15) R1: And they know carbon dioxide is gas, they absolutely know that.

--- 5 turns omitted

16) R1: How about they are elements?

17) VN: They are elements, that will work.

---- background chatter...

18) R1: Oh we can't B, we can't use atom .. We can use atom, I lied I lied I lied

19) R5: Oh it was there

20) R1: We can use parts of atoms.. we can't they we can't say mass number or atomic number.

21) R2: You can use the word atom?

22) R1: It says, um , its on (page) 20 in this one. (Reading from MDE 2000 Curriculum framework) Classify substances as elements compounds or mixtures and justify these classifications in terms of atoms molecules. Element compound mixture They are? Do these agree (pointing to page and asking R4) ..

23) R2: Elements

24) VN: Elements

25) All: Elements

During this discussion, R1 continues to consult the standards documents, finding that the standards do in fact indicate that Michigan expects that middle school students should know about atoms (turn 18). However, by that point, the committee decision process had begun to coalesce around changing the distracter to use the term “element” rather than the term “atom.” This conversation is indicative of boundary crossing, as the issues considered include those from the state and those for the student. Not all item discussions included these two elements, but most included more than one type of consideration, so the committee was not looking purely at wording or difficulty, but often considering the items from multiple perspectives.

In the next example, some of the challenges these committees face become even clearer. This example calls for some additional background, however.

Note that in the last example, R2 read part of the question and focused on the phrase "best describes." This phrase is significant because the students are given more than one correct choice and are being asked to choose the best one. The inclusion of more than one correct answer then allows the question's responses to be placed on a measurement scale, something important for the psychometric models common in educational measurement. One of the reasons so much time is spent in these committees on reviewing the choices is that it is important to have a number of student responses associated with each of the choices to show that the item is separating students based on a level of knowledge the students are theorized to have. This knowledge is identified in psychometrics as the "construct." Michigan's manager of MEAP testing, Joshua Martinique, discussed earlier in transcript 7.1 his interest in tests that measure multiple constructs. The questions these reviewers are considering, however, all map to a single construct.

a) balance b) digital scale c) spring scale

Photo 8.2 - Three types of measurement devices

The three devices shown in Photo 8.2 are used for measuring forces.[42] Photo 8.2.a shows a balance used to measure weight (the force of gravity). Many science classrooms have updated versions of these where the weights are attached and adjustable, called a beam balance. Photo 8.2.b is a digital scale that could be found in household kitchens, as well as in science classrooms. Photo 8.2.c shows is a spring scale that is hanging from a wooden beam. Of the three, the spring scale would be generally the least accurate, but also the least expensive. It is also commonly used in agricultural settings

[42] All three photos are from Wikipedia, accessed on February 8, 2008.

and has an important characteristic that the other two do not: it can measure force in any direction, whereas the other two only measure downward force from gravity (weight). While only two of the teachers I studied have a unit dedicated to scientific measurement, it is a fundamental concept that is part of the state, as well as national, science standards.

In Transcript 8.3, there is a discussion underway about a particular item that has been reviewed before. Note that several of the reviewers mention or begin to mention “spring scale” as they enter the conversation.

Transcript 8.3 - Content committee reviews a question dealing with the spring scale

---16 lines omitted

1) VN: OK what was this guy [item] flagged for?
2) R1: p-values
3) Others: p-values
4) VN: P-values that's the CL-sticker with the little correlation..
Option C also has a positive point biserial so there is a very attractive [option] C
What do you all think about the item?...
Why did they miss the item would be my question?
5) R4: The spr..
6) R1: Because
7) R5: Well, its spring scale
8) VN: Oh
9) R1: It's being taught incorrectly or not taught so they don't know and they guess.
Spring scale is not being taught.
10) VN: Should it be?
11) R1: Yes
12) R4: Yes
13) R1: Until we have new grade level content expectations
[inaudible]November

--- 16 lines omitted

14) BB: So what is the argument about spring scales
15) VN: Well OK Newton’s of force for the spring scale.. and that with the balance you are looking at the same thing you are looking at the force of gravity you just looking at it
16) BB: Using a little mass as uh calibrated a balance against
17) VN: Y’know I mean a good student could look at that and say well shoot I could use a spring scale to measure mass
18) BB: In some respects a spring scale could be more useful because you wouldn't have to carry those little weights around with you.

--- 8 lines omitted

19) R1: Well the reason the reason kids are picking C is that weight is a gravity issue.
and the spring scale operates on the fact that it is actively pulled on
That's exactly what is going on in the
20) VN: And the balance would be completely useless if you didn't have gravity

Transcript 8.3 - Content committee reviews a question dealing with the spring scale

21) R1: But um for finding mass you don't use a spring scale because it doesn't give you units in grams it gives you units in Newton's which is a force vector

22) VN: Which could be converted?

23) R1: It doesn't say converted, "Most useful" I like it I think its a good

24) VN: The most useful way of gettin mass cause gravity does work

25) R6: Balance is balance is is work

26) VN: Balance is the most useful way to get mass (the correct answer).

--- 7 lines omitted

27) R3: Its under Physical Sciences Matter and Energy (state standards section)

28) R1: In the using. OK. That's what I thought

29) R4: Concepts like: mass, weight, area, volume; array of measuring devices: metric ruler, graduated cylinder, balance, spring scale.

30) R6: Balance is commonly taught as a mass measurement.
So its probably still a good question.

31) R1: You know what's interesting is we list it as a measurement tool but the concept we list it being applicable to are the measurements of length mass weight density area volume and temperature not force
So are we leading teachers to believe .. that there's a huge gap there between
We say yeah you should teach your kids spring scale but it does not in any way relate to any of the components of the benchmark?...
So maybe our teachers are teaching it incorrectly because .. they're trying to figure out why spring scale is listed here to teach the concepts of length mass weight density area
And are they teaching that its a downward pull that is somehow related to weight.

32) R4: They must

33) R1: So now I have a problem with the benchmark listing spring scale and force isn't

34) R3: But weight is measurement

35) R1: It gets at weight and mass is not weight mass is another matter of an object leave it

36) R4: Keep in mind what they are asking for with these questions
The mass of a solid object

37) R1: That's a good question

38) R3: The most useful way

39) VN: The most useful way is a balance
and kids whether teachers are teaching it or not

40) R4: I think we should keep it

41) VN: It is most useful

42) R1: Yes, yes

In this discussion (which I have shortened for brevity) the panelists are discussing the most useful way of determining a solid object's mass. The phrase *most useful* is important, as is the very specific distinction between the concepts of weight and mass.[43]

[43] Mass is technically the amount of matter and weight is the effect of the matter being pulled by gravity. The mass of an object is constant, while the weight can change based on its location.

The issue with the item that this committee is wrestling with is that a high number of students chose the spring scale as the correct answer, where the correct answer was actually the balance. R1 makes the point that teachers are not properly teaching this concept to students (line 9). She also highlights the contradiction in the standards in that spring scale is listed, but not the concept of force (turns 27-33). The vendor then clarifies that the spring scale measures force (in *Newtons*) rather than weight. After some discussion about options, the group ultimately decides to keep the item as it is. The phrase “good question” that R1 utters (turn 37) was a way that I observed this group would often move towards a decision to keep an item. The consensus is described by the vendor representative (turn 39), who restates the most useful way to obtain mass is through a balance, even though a large number of students chose the spring scale option.

Before turning attention to reviews of MEAP results in the schools, it is important to observe that not all schools in the state are identically equipped, and it is quite conceivable that the measurement devices available across the state differ from school to school. Further, it is conceivable that some students, perhaps those in rural areas, might encounter devices such as spring scales in their out-of-school activities while other students, perhaps in more urban areas, might not. In attempting to create a realistic context for the question with answers that map onto a scale, the phrase *most useful* can be seen as a value-laden cue where issues of government standards, classroom teaching, school equipment, financial connections, and local culture are all at play in what Swales (1998) refers to as the microtext of a test item.

Boundary Practices in Schools: Case Highlights from Swallow and Avon Falls

The MEAP tests were administered during October for the 2006-2007 school year. The results were then released on January 9, 2007. The work the state was doing in the interim included the scoring of constructed response items, waiting for a contractor to manually evaluate student responses, and statistical work necessary for properly reporting the results. Because of the contingencies of this work, the exact timing of the release of the results is not generally known until the state announces the scores are ready. Once the state release of scores to the schools occurs, then each school can initiate its own process of distributing and making sense of the results. Within this schedule is a window of opportunity for a researcher to observe the school's interpretation process. Being able to observe this process requires being able to visit the school at the specific time of the review, which will likely occur on a specific day after the results are released and when students are not in the school. The school holiday and professional development calendar then become important factors in observation of school MEAP result review events.

Because of these schedule constraints, I am only able to present Swallow and Avon Falls' MEAP interpretation processes. It was not possible to observe the comparable work in Dorchester and Crimson. Hardy was using a data warehouse for the first time during the 2007-2008 school year to support MEAP interpretation, and I will discuss my observation of that process in the next chapter.

Case Spotlight: Cross-Grade Work in Swallow

At Swallow Middle School the building principal, Burt Wainwright (BW), played a leadership role in the interpretation of the MEAP results. That year, he was also the curriculum specialist in the district, which included an elementary school in addition to the combined secondary school. He was the one who logged onto the secure OEAA website and downloaded scores. He then distributed the results to teachers and prepared presentations for the board of education. I interviewed him about this process and received examples of the artifacts he created, including the memorandum he sent to his teachers summarizing the results and the graphs and displays he created to use in making presentations. I was able to document the existence of at least eight steps in Swallow's MEAP interpretation process, as illustrated in Figure 8.1.

<u>Structure of Swallow's Interpretation Process</u>

While there was both interview and documentary evidence of this process at Swallow, I was only able to observe parts of it. Table 8.1 summarizes the steps illustrated in Figure 8.3 and the evidence that was collected. Note that for the departmental review of the science scores (Step G), I was only able to collect video without audio.

Table 8.2 - Major steps in Swallow MEAP interpretation process.

Step	Controlled by	Date occurring	Evidence
A.State releases preliminary results	State (MDE)	1-9-07	Fieldnotes, photos, public documents
B.Memo to staff, initial spread sheets & charts	Local Calendar (BW)	1-17-07	Collected memo
C.State released final results	State (MDE)	1-22-07	Public documents
D.Additional charts developed	Local Calendar (BW)	1-17-07 thru 2-4-07	BW's PowerPoint from K-12 meeting
E. Board Presentation	Local Calendar	2-4-07	none
F. Results sent to staff	Local Calendar	1-17-07	none
G. Department Review	Local Calendar	1-30-07 (planned) 2-7-07 (occurred)	Video, no audio
H. K-12 Meeting	Local Calendar	3-13-07	Fieldnotes, audio, and documents

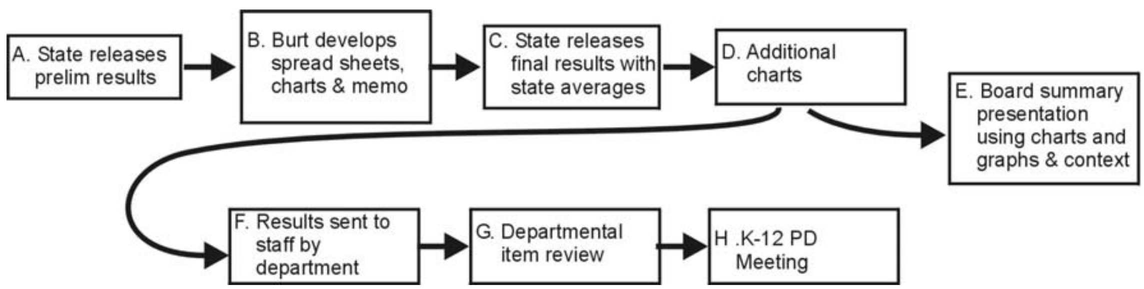


Figure 8.1 - Swallow Schools MEAP Interpretation process

The documents presented in Document Set 8.2 show some of the highlights of the MEAP interpretation process at Swallow. Document 8.2.a is a close-up of the state MEAP report for Swallow Middle School, showing student performance on various items that are grouped into strands. The strand is a superordinate category for state learning objectives. These objectives are also known as grade level content expectations (GLCEs, pronounced "glicks"). This report shows the specific questions that were released, the GLCEs that the questions related to, and then details about the percent of the students that chose various answers. Other state reports, including the Item Analysis Report, show the correct choice and the number of students that selected it. Document 8.2.b is the memo from Burt to his staff introducing the results. He told me that he always prepares what he called "a narrative" like this to present the results. I also saw and photographed a similar memo for the 2005-2006 school year. Note that in the last paragraph he is describing the meeting schedule (Steps F and G described in Table 8.1 and illustrated in Figure 8.1).

Document 8.2.c in this set is one slide from Burt's presentation to the staff in the K-12 meeting (Step H). He indicated that this presentation was the same that he had given to the board (Step E) as well. The mastery indicated in this chart comes from the overall scores as defined by proficiency level. This is calculated using the specific items chosen by students on the MEAP, as well as scores given for constructed response items.

The K-12 meeting (Step H) this discussion highlights occurred in the school's media center or library. Burt Wainwright led the discussion and began with a Powerpoint presentation where he summarized the results by grade. This presentation included the same types of visual displays I had seen in his office in January. Early in the presentation he showed Swallow's overall results in comparison to schools in their athletic conference. He explained this comparison for me in our January meeting. Even though those schools were not within the same county or ISD, they were matched in terms of size and resources, so they made a good comparison group for reporting school performance. After his overview of school performance, Burt and the elementary school principal divided the teachers into teams based on subject area. They instructed the teams that they should review the subject area MEAP results and be prepared to present to the group a summary of their plans for improvement.

a) Close-up of portion of Swallow Middle School's Science MEAP report

No. of Students Assessed = 61

RELEASED MULTIPLE CHOICE									
				PERCENT RESPONDING					
STRAND Domain	Released Item Number	Benchmark Code	Item Type	A %	B %	C %	D %	Omit %	Multi %
CONSTR. KNOWLEDGE									
Constr. Knowledge	25	I.1.m.2	Core	8	25	52+	15	0	0
Constr. Knowledge	47	I.1.m.2	Core	20	8	62+	8	2	0
Constr. Knowledge	23	I.1.m.3	Core	10	61+	20	10	0	0
Constr. Knowledge	24	I.1.m.3	Core	44	31+	16	8	0	0
Constr. Knowledge	30	I.1.m.3	Core	13	33	26	28+	0	0
Constr. Knowledge	3	I.1.m.5	Core	5	8	82+	5	0	0
Constr. Knowledge	5	I.1.m.5	Core	10	20	28	43+	0	0
REFLECT KNOWLEDGE									
Reflect Knowledge	8	II.1.m.1	Core	16	51+	18	15	0	0
Reflect Knowledge	21	II.1.m.1	Core	25	18	18	39+	0	0
Reflect Knowledge	7	II.1.m.5	Core	11	15	64+	10	0	0
Reflect Knowledge	9	II.1.m.5	Core	7	70+	3	20	0	0
Reflect Knowledge	31	II.1.m.5	Core	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

STRAND Domain
Ecosystems
PHYSICAL SCIENCE
Matter & Energy
Matter & Energy
Matter & Energy
Matter & Energy
Matter & Energy
Motion of Objects
Motion of Objects
Motion of Objects
Waves & Vibrations
Waves & Vibrations
Waves & Vibrations

b) Burt's memo presenting MEAP results/plan

To: MS/HS Staff
Re: Preliminary Fall 2006 MEAP Results
Date: 17 January 2007

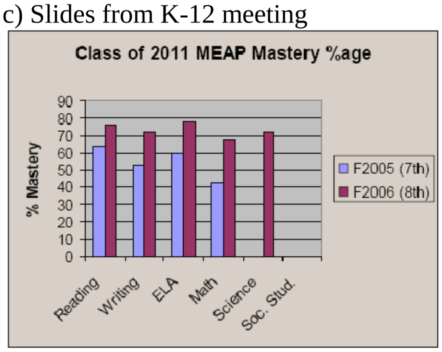
Our preliminary MEAP scores from the fall testing for grades 3-9 were released at the end of last week (with accompanying server issues at the state level—it's comforting to know that others have computer issues as well!). Our overall results are displayed on the following sheets.

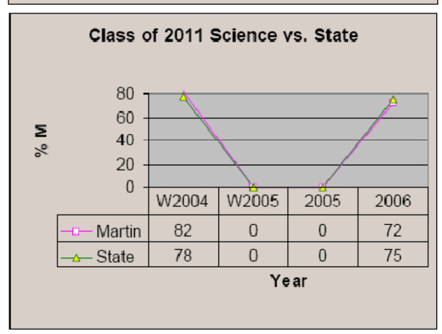
In general, I think the results look very encouraging! A word of caution, however--at this time we have not been given overall information on how we relate to the state average. As soon as this information becomes available, I will share it with you. It is possible that scores are up across the state and our gains are therefore not as significant. We simply will not know until state averages are released.

Last year we had some scores that fell below the Middle School AYP threshold for ELA and Math. That level remains at 43% again this year, and I am happy to report that all of our scores surpassed that mark for the 2006 testing cycle! This will help tremendously to balance out those scores that fell below the line last year. Looking ahead, the next increase in AYP threshold will come with testing next fall when the cut-off will move to 54% (if there are no changes in the NCLB legislation between now and then). The good news is that we are already there in ELA, while Math scores showed gains.

I'd like to applaud all of you for your efforts! We will discuss this more at our March in-service and we will also be scheduling department release-time work sessions in the near future. Prior to that, department chairs will be sharing released items from the test with you, along with detailed information about where our students struggled. Hopefully, we can continue to make adjustments necessary so that our performance will be even better next time around!

c) Slides from K-12 meeting





Document set 8.2 - Selection of Burt Wainwright's documents related to MEAP results

<u>Reviewing Items in Swallow</u>

The science team that reviewed the MEAP results included eight teachers. Among them were Paul Bond (PB), the science department chair Toni Donnard (TD), and the sixth grade science teacher Cathy Amazingly (CA). Others from the elementary school and the physical education teacher were part of the group, but were not as engaged in the review of the eighth grade results as those three just named. In transcripts of this meeting I refer to these other teachers as (T1 and T2). The discussion of specific results lasted about 60 minutes. During this time a total of 13 assessment items were reviewed. The process of this group meeting began with a discussion of what the criteria should be for selecting an item to review and then item reviews similar to the one shown in Transcript 8.4:

Transcript 8.4 - An item is reviewed in Swallow	
1)	TD: We had a problem with number 29 They didn't know which way..what is a South wind what is a North wind what is an East wind
2)	CA: You're gonna see a difference this year with Mr. Wylke (resource provided by Kellogg Foundation grant) is doing now setting up the uh weather station um The year that I did that ..in the classroom I did one in 5th grade they scored pretty high in the MEAP test because there's like six questions on the weather.
3)	TD: OK
4)	CA in the 5th grade science we're.. he's building that science thing the science weather station and we have that thing from you and my class made weather vanes.
5)	TD: So hopefully they'll know what wind direction is…
6)	CA: They didn't know what a rain gauge us. They called it a goage and they didn't know what a goage was. I'm not kidding. Every single kid.
7)	TD: A little French perhaps? goage
8)	CA: It killed us. It killed us last year.
9)	TD: Well understanding vocabulary the names of the equipment is …vocabulary there.
10)	CA: Six questions on weather. But next year the State may decide not to focus on those GLCEs because they pick and choose.

The number 29 referred to in Transcript 8.4 (turn 1) is the same number as would be shown on the MEAP item report (Document 8.2.a). This number is the number of the item from the MEAP test. Because images of the test items were released to the public along with the results, the team of teachers was able to review both the original item and details as to which students chose what answer on the state’s standard MEAP reports. I have included that item and two others in Document Set 8.3.

Just as the review of items under development for the MDE was analyzed and coded for topics using the qualitative coding structure described in Chapter 4, all of the item reviews in the school were coded using the same coding scheme. Item 29’s review (Transcript 8.4) was coded for three topics: classroom equipment, terminology, and a cross-grade discussion. Half of the item discussions for this group included some cross-grade issues. Another example of both cross-grade and instructional alignment issues is in Transcript 8.5, where student performance on question 33, involving the nature of matter, is discussed. This discussion leads to an important diagnosis at the end about what is being taught in the classroom (turns 16-21).

Transcript 8.5 - Item number 33 is reviewed by the Swallow K-12 science team.

1) TD: OK what is formed when two hydrogen atoms combine with one
2) CA: No they've got to have that.
3) TD: one oxygen atom. Two atoms combine with one oxygen atom. no.. . when two oxygen atoms are combined with one hydrogen atom. Do you get a mixture a mineral an element or a compound?
4) CA: Do you talk about that in 5th grade?
5) T2: We have one lesson on that
6) T1: That's a lesson in the eighth grade isn't it?
7) CA: Before we didn't really do it in sixth but now with this new program so

--- pause and conversation

8) TD: Well what percentage did we mess up on that one?
9) CA: 48 percent said B, 30 percent said D (the correct answer)
10) TD: 48 percent said it was an element. Number 33 they said it was an element and not a compound.

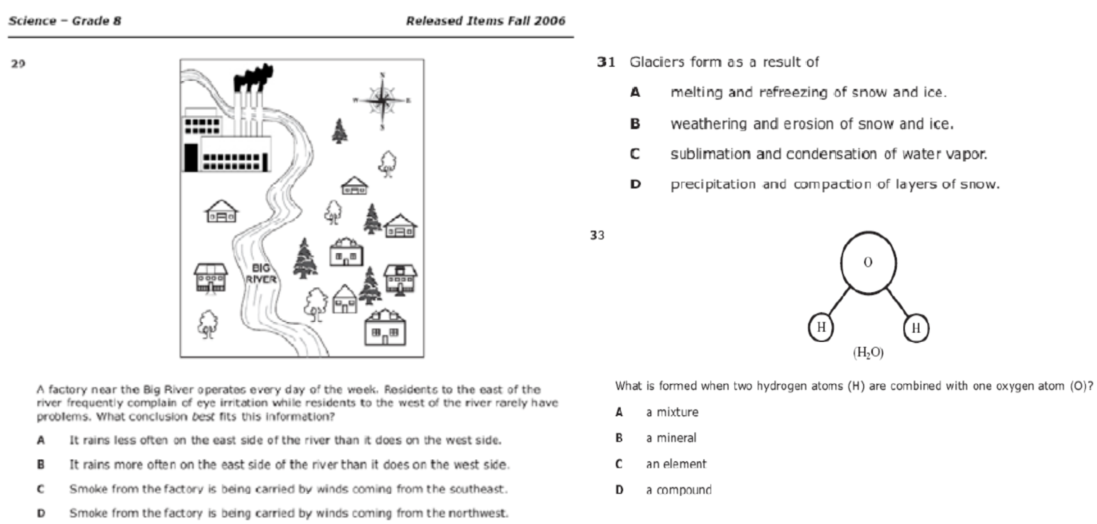
--- pause

11) T1: So what was the percentage again?
12) TD: 43 percent said it was an element.

--- pause

13) CA: I mean you could spend a lot of time on--
14) TD: Yeah I know but this is where we are having.. you can see where
15) T1: This gets introduced in 5th grade. We don't do anything more than know that there's hydrogen and oxygen
16) TD: Do you (turning to Paul) do much with elements and compounds? You do some in
17) PB: Eighth grade a lot--
18) CA: A lot of the stuff we're struggling with?
19) TD: Sometimes, sometimes not?
20) PB: Usually not.
21) TD: Usually not.
22) CA: Now it's being introduced now it's in sixth grade.
23) TD: It is in sixth grade?
24) CA: It is now. That's why I said not to spend a lot of time on it.

This discussion around Item 33, a nature of matter question, is important because it links to the study data presented in Chapter 5. In Transcript 8.5, Cathy Amazingly (CA) and Toni Donnard (TD) lead the discussion. The concept that this question covers (shown in Document 8.2.c) is central to understanding how to describe and categorize matter. It is in the same conceptual area as the discussion in the state's review of the question involving atoms and elements shown in Transcript 8.2. In this discussion, a teacher (unidentifiable on the audio file) asks in which grade the topics of the composition of matter into elements, compounds, and mixtures are taught. According to their discussion, only 30% of their students chose the right answer.[44] Then at the end of the discussion, Toni Donnard asks Paul Bond how he covers the subject. He initially says (turn 17) "in Eighth grade a lot." If he did cover it in eighth grade, rather than seventh, the students would probably not have mastered the concepts unless they were taught early in the year, as testing occurred in October. However, on her challenge (turn 19) of "sometimes, sometimes not?," Paul admits he may not cover this material in depth, when he says (turn 20) "usually not."


Science – Grade 8 | Released Items Fall 2006

29

A factory near the Big River operates every day of the week. Residents to the east of the river frequently complain of eye irritation while residents to the west of the river rarely have problems. What conclusion *best* fits this information?

A It rains less often on the east side of the river than it does on the west side.
B It rains more often on the east side of the river than it does on the west side.
C Smoke from the factory is being carried by winds coming from the southeast.
D Smoke from the factory is being carried by winds coming from the northwest.

31 Glaciers form as a result of

A melting and refreezing of snow and ice.
B weathering and erosion of snow and ice.
C sublimation and condensation of water vapor.
D precipitation and compaction of layers of snow.

33

What is formed when two hydrogen atoms (H) are combined with one oxygen atom (O)?

A a mixture
B a mineral
C an element
D a compound

Document set 8.3 - Fall 2006 Science MEAP Released Items: 29, 31, and 33.

Since Paul was a case study teacher, some details about his teaching were collected. The first photo from Chapter 5 shows a Periodic Table of the Elements on the back wall of Paul's classroom. The element descriptions are made with yarn and beans. A review of figure 5.8 at the end of Chapter 5 shows that chemistry is the last subject Paul teaches in seventh grade. Given the common reports by teachers that seventh graders do their most thoughtful work in the middle of the year, it is possible to see this craft work as the kind of end of year activity that does not demand that much of students. With this explanation, Paul can be seen as a teacher who is not prioritizing this key state

[44] There is some discrepancy between the option these teachers are discussing and the item itself. Either 40% of the students chose option b (mineral) or when the teachers are discussing the question they mean option c when they say option b.

learning objective in the way he sets up his curriculum. He puts the nature of matter at the end of the year rather than earlier, when teachers like Betsy Dearing indicate that students do their best work. An alternative explanation is that since Paul teaches both seventh and eighth grade science in this small school, his placement of chemistry at the end of seventh grade is designed to prepare the students to engage it again in the eighth.

Summarizing the Group Discussion

Transcript 8.6 - An item is reviewed in Swallow	
1)	T3: OK some of the high high points that .. we saw that could be a uh uh a fix K through 12 these are the key things that we thought that if everybody work on them it would help One is volume mass density and weight and knowing the difference between those things these seem to be a big uh problem from in high-school um Different kinds of instruments like weather instruments. Also science tools, graduated cylinders and what's the one thing you said?
2)	T1: Anemometer ?
3)	T3:Anything any kind of science tools, uh spring scales, they just need to be real real familiar with those Another thing because science crosses math's path a lot, the kids have got to know their math facts. They are NOT retaining them. And we've got just continually spiral and do that Any kind of mental math cause they're lost, their stuck. How can they do these big science formulas when they don't know what 2 times 3 is without counting on their fingers?
4)	T1:Have them do averages in all the classes, averages or scores or something
5)	T1: Averages is also a big deficit. OK. And I think that's just about it.

In their report to the group, the team summarizes their discussions, which included topics of teaching students about mass, density/volume, and equipment such an anemometers and spring scales. The anemometers measure wind velocity. Interestingly, spring scales did not appear in any of the item discussions I recorded at Swallow, nor did they appear on any of the MEAP items for that school year, but they were mentioned by the science team in their report. Note that in the discussion of the spring scale item from the state review (Transcript 8.3), several teachers seemed to spontaneously identify the spring scale as the root issue with the question. This suggests that issues of measurement equipment in the classroom and their relationship to the scenarios used in high-stakes science test items may be an enduring challenge for some educators and part of their knowledge of practice (Lampert, 2001).

Avon Falls MEAP Review

In another part of the state, Avon Falls personnel followed a process similar to Swallow's as they interpreted their 2006 MEAP results in early 2007. Avon Falls is a larger district than Swallow, with a dedicated curriculum director, Janey Fess, who performed the function of downloading and disseminating the results that Burt Wainwright performed at Swallow. She first sent the reports to the building principals. According to Bob EnSpania, the district superintendent, the principals were to share the reports with the staff and then make individual presentations to their board of education. Janey herself coordinated a science review that included representatives from the different schools followed by meetings with individual departments.

This first meeting Janey held to focus on science was a cross-grade (K-12) event that preceded the departmental review, which is the opposite order from what occurred at Swallow. But the order was influenced by the weather and may represent a matter of scheduling convenience rather than a formalized process. I was originally invited by Janey to attend the department meeting a few days after the state released the results. Janey suggested I could observe the teachers (who I had not yet met) as they reviewed their scores. After discussing the idea of surprising the teachers at this meeting with their principal, Stan Dubovski, I decided not to attend that meeting because the teachers had not yet consented to being part of the study. Stan agreed with me not to surprise the teachers in this way. As it turned out, an ice storm caused the original meeting to be cancelled, and by the time the rescheduled department meeting occurred later in February I had already met the Avon Falls teachers, who then gave me permission to observe their meeting.

Since the K-12 meeting had already occurred, I spoke with Janey Fess about what she covered when she gathered the science leaders from the different schools together. Transcript 8.5 shows a portion of that conversation related to how she perceived the teachers' response to interpreting their MEAP scores. The teachers she is specifically referring to in this transcript are the science leaders. However, in other comments, she expressed a similar view for teachers across the district who are attempting to use data to inform their instruction.

Transcript 8.7 - Janey Fess Describes the Avon Falls K-12 science meeting.

1) JF: I think it's the learning curve. What I am finding is that teachers don't have the knowledge as to what to do next.
2) PP: uhum
3) JF: And let me talk about um the process I used in working with the science teachers and what happened as a result of that
4) I had I divided the uh MEAP information, the item analysis and the school summary as well as giving them a copy of the the released items.
5) PP:uhhuh
6) JF: And they worked in I had them grouped of two teachers that looked at 5-6 from a 5-6 perspective. Actually I should take that back from a 5-8 perspective looking at the 8h grade test.
7) PP: uhhmm
8) JF: And had a high school teacher that worked with them.
9) PP: When you say the released items you mean this year's items.
10) JF: Yes, this year's items and we looked at that. And then I had an elementary group .. who looked at the test items. And I asked them to go through and look at what our strengths were and our opportunities were. And .. to start and to use the test questions to sort of think about where the kids were at. And so they were talking within in their separate groups there were two groups .. there and looking at the test and looking at the released items and then we came back together as a group to see what we could learn from looking at those grade levels.

Without textual or observational evidence of this meeting, I cross-checked Janey's account with Betsy Dearing a few weeks later. Betsy's version aligned with Janey's, and she mentioned the issue of content alignment in reference to the MEAP results showing poor performance on a question about glaciers. This "Glacier Question" turned out to be a useful conversation starter with other teachers who had reviewed their results, as many schools did poorly on it.

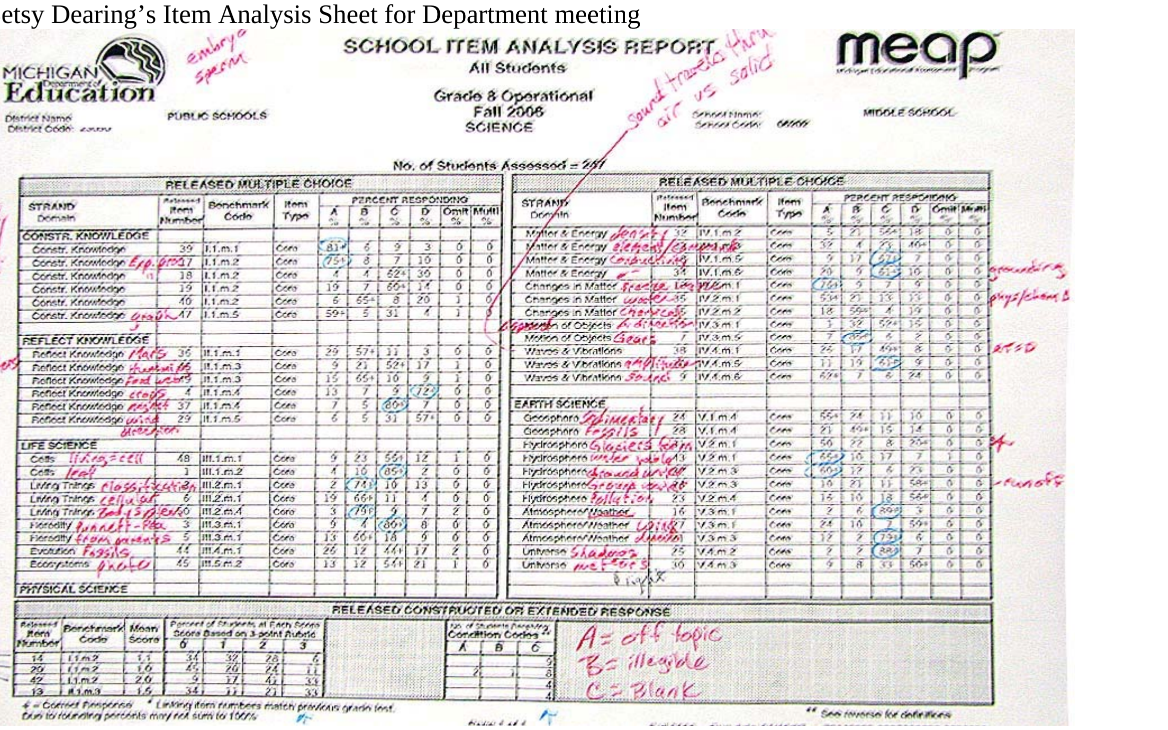
When the department review occurred, I observed and took field notes, because the science teacher Jim Heinrich had asked not to be taped. This departmental review had a similar discourse structure to the K-12 review at Swallow and the state meetings. After an initial orienting discussion, the meeting proceeded on an item-by-item basis. My notes record 22 MEAP items were reviewed, and there were several other discussions related to the review process. The first discussion was an orientation to the reports and some conversation about the fairness of testing. Many of the item-oriented discussions involved issues relating to classroom equipment or local teaching methods contrasted with the approach represented in the MEAP. After the departmental meeting, when Jim

Heinrich had excused himself the other two members, Betsy and Jonathan Brunson, allowed me to audio record a discussion we had about the meeting.

Transcript 8.8 - Betsy Dearing's account of the Avon Falls K-12 science meeting (from an email).
The meeting isn't so fresh in my mind any more, but I can tell you what I got out of it. We spent most of the morning going over the MEAP results. I worked with a teacher from the 5/6 building and we looked at the 8th grade MEAP test results. Elementary teachers looked at the 5th grade MEAP results. At the end, I felt like some of my test questions do fit with the MEAP: the questions that I copied from other textbooks that use data and have a couple of multiple choice questions and some of the questions I have created that require students to read data. Also, we looked at extended response questions, their rubrics, and some sample writing. Again, I felt that some of my test questions were asking similar style questions (specifically, on the digestive system test when I ask students to pick and explain which is the healthiest or least healthiest). Again, there are things the students are consistently not getting correct and I will try to emphasize those items more in the future or do additional activities to help the students learn the concepts. Some items we were very confused about why the students didn't correctly answer the questions and we weren't sure how to address those concepts. One question about glaciers made us realize that although we talk about glaciers, we don't ever talk about how they are formed. (More students got this wrong and chose one wrong answer than those that chose the correct answer.)

Just as with Swallow's meeting sequence, the documentary evidence combined with the personal accounts shows activity relationships for Avon Falls' interpretation of its science MEAP results. While the specific steps were different, Avon Falls' process, like Swallow's, was largely structured by the item report. Betsy's marked-up item report (Document 8.5.a) shows the work she did at the K-12 meeting, and she used that as a guide for the departmental review. I was able to confirm, as with Swallow, that the process had occurred the previous year in a similar manner when Betsy showed me her notes from the previous year (Document 8.5.b). She commented that they went through the same process as last year that included identifying specific vocabulary words to work on. But, she said, she wasn't sure if they had followed through on that plan. Similar to the reflection of Pete Darmond; who near the end of Transcript 6.1 expressed doubts about the extent to which districts are able to actually translate the use of assessment information into changes in classroom practice, Betsy admitted that the review process had not translated into any action in terms of classroom practice. None of the Avon Falls science team item review discussions I observed included any mention of the prior year's review or any efforts the teachers had made to adjust their practice.

a) Betsy Dearing's Item Analysis Sheet for Department meeting



b) 2006-2006 Review summary sheet

Science MEAP

Strengths
- no multi answered questions
- very few omitted questions

All - using/evaluating sources of information
All - awareness & sensitivity to natural world
7th - Plants make & store food
7th - How systems & processes work together in animals
5th - Patterns of relationships (pred/prey - ecosyst.)
6th - Use forces to move objects

Weaknesses
All - using scientific tools
All - evaluating strengths & weaknesses of claims or data
All - awareness & sensitivity to natural world
7th - contributions from heredity & environment
5th - population dynamics
6th - elements, compounds, mixtures
6th - unbalanced forces causes motion causes motion
7th - How light is required to see objects
7th - How light interacts with objects
5th - Rock formation

c) 2006-2007 Department meeting summary

Science – PD MEAP Review February 16, 2007
7th and 8th grade

Weaknesses
- **Reading and interpreting data**
- **Experimental design, especially variables**
- Not enough time to cover all benchmark information
- Formation of glaciers (more students selected one incorrect answer)
- Reading a diagram of the "fossil record" to predict climate change (unrealistic fossil record)
- Writing to the facts provided from data

Strengths
- **Human body (a question was finally on the MEAP!)**
- **Atmosphere and weather**
- **The more time spent on a topic, the better student results are (human body and weather are topics given a lot of time).**
- Forecasting weather (good illustration provided on the MEAP)
- Shadows (good illustration provided on the MEAP)

Strategies
- Use more illustrations and data tables
- Conducting investigations was emphasized on the MEAP. We would like to increase our classroom emphasis on conducting investigations, but we are limited by a lack of materials and time to allow for the increase of experiments which would provide students increased exposure and practice.
  - Not all experiments result in data collection, but do have variables. (Experiments would focus on data collection and interpretation or variables.)

Document set 8.4 - Documents from Avon Falls MEAP results review

Analysis of Topical Clustering in Item Discussions

This boundary practice analysis focuses on discussions related to MEAP items across the two main contexts. It draws on the three item discussions observed at the state and the sets of reviews of results from two schools. Since the same qualitative coding scheme was used for all these events, the content of the discussions can be compared. Table 8.2 summarizes the number of times various qualitative codes[45] were assigned across the five events where the MEAP items were discussed, categorizing the qualitative codes into four groups: state issues, local and school issues, students, and other.

Table 8.3 - Summary of issues raised in MEAP item reviews

		State MDE			*Schools*	
	Number of topical occurrences **(items)**	**Bias & Sensitivity (14)**	**Stat Review (10)**	**Content Review (38)**	**Swallow K-12 (13)**	**Avon Falls Dept (22)**
State Issues	Graphics in MEAP items	9	4	12		6
	Terminology in MEAP items	4	2	10	2	4
	Standards and benchmarks	1	2	12	1	2
Local and School Issues	Cross domain connections			1		
	Cross-grade meetings/process				6	2
	Inquiry methods					2
	Alignment between instruction and standards			2	1	
	Classroom equipment		1	2	4	2
	Learning contexts		2	6		5
	Teaching methods		2	1		
	Vocabulary in the classroom			1	2	1
Students	Misconceptions		2	1	2	
	Student ability		1		6	
	Student cogitative processes		2	2	2	3
	Student test taking skills					4
	Practical knowledge					1
Other	Historical Change				2	
	Critical about assessments					3
	Policy: accountability		1	1	1	
	Policy :special education				1	2

Because the overall sample is small and singly rated, the information in Table 8.2 is used to illustrate some possible patterns these meetings exhibit rather than to generally

[45] Each of the topics listed in the second column had a specific code associated with it or is the combination of several individual codes. The specific codes (ex: AAS-G for assessment artifact at the state level, sub-code of graphics) have been omitted for clarity.

describe the relationships between state and school science assessment systems in Michigan, as a larger sample might.

One pattern this analysis illustrates is that in both the state and school meetings, there are a mixture of topics related to both contexts. These data suggests that both types of events exhibit boundary properties. Some of the more common topics in these discussions are those dealing with cognition, standards, terminology, and classroom equipment. It is unclear from this small sample whether these two types of boundary discussions are complementary, for example if with more meeting observations if meetings in one location would be more dominated by certain types of topics than in other locations or whether the composition of the groups in terms of participants' backgrounds influences what topics are discussed. It is also unclear the extent to which the science sub-domain is a factor in what is discussed. Do the physical sciences tend to draw out questions of equipment more than, for example, life sciences? These are questions for more thorough analyses with a larger data set.

Across these five conversations is a possible asymmetry in topics that is worth attention. For example, while the graphical representation of the MEAP items was frequently discussed in the reviews at the state and in one school, no discussions of the graphics that students might encounter in classrooms from their textbooks, assessments, or teacher classroom illustrations were observed in any of these events. While Chapter 5 illustrated the wide range of symbolic competencies students across the seven classrooms were asked to master, none of the discussions related to the MEAP items (at either school or state) considered how the tasks students encounter in their classrooms and the MEAP test might be aligned. Other issues related to classrooms, including the natural contexts of student learning or teaching methods, do receive attention, but not the multimodal communication (Lemke, 2002) that students may experience in their everyday classrooms or how these multimodal experiences relate to the accountability test.

The individual MEAP item is clearly important in describing the systemic relationship between the state and teacher assessment systems in Michigan. In all five of the review events, the order of items was not based on standard or strand. Crimson and Dorchester's teachers all reported similar item-based review processes, although they were not studied. Hardy Middle School is a little different, as the year that I observed them, they were beginning to use the MI-Tracker data warehouse system discussed in Chapter 7. But, as Document 7.2 showed, Dianne VanderMiller used specific MEAP items to structure her work with the science teachers. The MEAP item can be thought of as central to mediational artifacts, a type of boundary object (Star and Griesemer, 1989) that connects the state to the school.

Beyond the Item: Broadening the Perspective of State-School Interactions

Because of its structuring role, the MEAP item is a candidate for being a boundary object. However, it is only one element in a complex of processes as it crosses boundaries (Moss, Girard, and Haniford, 2006). In the following three analyses, I will explore data in this study that support considering inter-systemic relations associated with the MEAP and school practices beyond the item.

The first of these analyses looks at leadership response. As the literature reviewed in Chapter 2 indicates, leadership plays an important role in the use of information in educational settings, and this study has yielded a glimpse of two different leadership approaches. The second analysis looks at macro-temporal structuring between the state and the school. In elaborating on the model of timescale and ecosocial level articulated by Lemke (2000), I find that the state MEAP process seems to influence a topology in school attitudes. The third analysis highlights the role of the individuals who pass between school and state and help to develop the MEAP items. These individuals, like Christy Connolly, could be called brokers (Wenger, 1998) or boundary spanners (Tushman and Scanlan, 1981).

Highlights of Leadership: The Basketball Conference and Manly Men Park

While the discussions related to the MEAP in both Avon Falls and Swallow highlight the issues of instructional alignment, there are also cultural and leadership dimensions associated with a school's interpretation of the state test. The two schools whose MEAP review processes I sketched above were different in a number of ways and had different kinds of principals. Swallow had one principal, Burt Wainwright, for both the middle and high school portions of a combined secondary school. Burt presented himself formally and served as the school's basketball coach. Avon Falls was a larger district and had a school for seventh and eighth grades. In Avon Falls, the middle school principal, Stan Dubovski, was comparatively informal. Since one of the themes raised by the literature reviewed in Chapter 2 is the critical role of leaders in information and data processes in schools, it seems appropriate to consider briefly how these two men approached being leaders of their schools' responses to accountability testing.

Transcript 8.9 - Burt Wainwright addresses the K-12 faculty about the MEAP results review
To begin though I wanted to quickly run and I know you have all seen it before
but I want to show you some of you from the elementary have not seen this the results from last fall. I actually used this at the board meeting I used this in the February 4th board meeting I recall and I wanted to show you what I showed them. I'll give you the little spin on it that we used. But we know that we did uh .. better we did much better this fall overall .
~ a few minutes of presentations (documents in 8.2.c part of this presentation)
All right and then in eight grade, these look like huge gains across the board and they were. OK, our eighth graders really did a nice job .. uh .. this fall. See there that's swing of 20-21 I guess percentage points in reading, so that's fantastic. In writing a similar sort of thing. Not quite as much but almost. ELA a big jump, 20 percentage points again. There's math and then their science score again. It's hard with science because they're only tested every so often, but they are right near that state average as well.
~ several minutes later after presenting grade by grade the results.
Now we're going to have you take that stuff and put it together on K-12 ticket. OK?
Talk with each other, look again at strengths and weaknesses. See if like we did a year ago if we can really identify K12 weaknesses. What steps we need to be doing. Sometimes y'know in upper El once those kids leave, y'know they're still your kids in a way you you kindof don't keep as close a track of them cause they're kindof off to the next area. And um for us in the secondary end we have to uh we're trying to find out about those kids coming .. what their strengths and weaknesses are .. and the coordination between our two halves is something I think it real important. And it is something we have talked about that with a school our size .. one of the good things with our size being smaller is we should be able to have uhm some quicker movement on some of these systemic issues because we're .. smaller.. we can just within the few people that are involved talk things out and get some changes made and so that is what we are looking to do.. OK?

At Swallow, But Wainwright was tall and commanding. He was present at the departmental review and led the K-12 review process. In every discussion I had with him about the MEAP, he took a balanced tone, noting the need for accountability but often describing its difficulties in practice as evidenced in Transcript 7.2. In the K-12 meeting, he presented the results to the district teachers using a slide for each grade and presented an overarching narrative of Swallow's improving scores and the need to shore up weak spots. Transcript 8.9 is excerpted from this presentation as he introduced the MEAP results to the collected faculty. After the results ware presented, he framed the work the teachers should do in teams by taking a cross-grade problem-solving approach.

The approach that Burt takes in this presentation is to address explicitly the school's response process by directing the staff across the two buildings to engage in data-informed problem solving: to think of themselves as part of one team that is responsible for the students across the grades. As the review portion of the meeting concluded, Burt and the principal of the elementary school, Jane Baily, led the teams in reporting their results. Burt, as the facilitator, called on each team to describe what they had found and their plans for improvement. The science team's report was shown above in Transcript 8.6. Burt's response to the MEAP could be considered explicit and direct.

The leadership approach in Avon Falls could not be more different. I also observed Stan Dubovski, but did not record him, over several parts of the day when his department met as part of an all-faculty professional development day in his middle school. During this event, in the school's library when presentations were being made and just afterwards at a faculty lunch where one group of teachers (on this day the science team) prepared a common meal, the subject of the MEAP tests did come up. In both parts of the meeting, Stan identified the general task of reviewing MEAP results, but did not present any results or describe the nature of the work the teams would do. In my interviews with him, he did not describe any of the work he did associate to his school's MEAP results nor show me any reports about the school, as Burt frequently had done.

However, the MEAP did come up in another way that shows a dimension of Stan's leadership response to it. While interviewing Stan, the story of an event related to the stress of taking the MEAP emerged. In this story, shown in Transcript 8.10, he describes the team-building process that began as a response to the stress of the time of year that the MEAP test is given.

Transcript 8.10 - Stan Dubovski and Manly Men Park	
1)	PP: Um I wanted to ask you um I parked out there next to Manly Men Park..what is that?
2)	D: I am so proud of Manly Men Park. Last year, it was just before MEAP, we were gearing up for MEAP And we had um everybody was kind of stressed out.

Transcript 8.10 - Stan Dubovski and Manly Men Park

So I got together all the men .. of the staff and I said Lets have a tailgate party put on just by the men.
We'll buy food and we'll cook food outside we'll tailgate for the women.
So we did that and we were we became the manly men.
That December, as kindof a ... staff morale builder I said let's do a Christmas dinner for the women and let's make a calendar.
We'll make a manly men calendar.

3) PP: Oh you mean with the photos and all that?

4) SD: So the front picture was all the men of the Middle School and dressed all like men in black or blues brothers..

5) PP: uh huh

6) SD: I have to show it to you.

7) PP: I want to see it

8) SD: You hit you hit a chord there and we ... and we did a big Christmas dinner for the women of the staff and we did one other big thing and of course they're going to reciprocate and they're better cooks and we get more food and there's more of them.
And we did the same thing we did the annual tail gate thing this year again and one of the teachers I was so proud of her said, not said snuck around got all the women teachers and said lets get together and lets name this…
And jokingly, during one of our tailgates because we had it in that area back there and jokingly I said we should name that area, before it was just an place by the dumpster, but with the new construction and everything it got landscaped real nicely I said it should be named manly men park I said I said that.
So in her memo to everybody she said lets make them a sign that says Manly Men Park.
I was so proud of her.
This was a young teacher 27, 28 To take kindof take that risk and to be silly about it the whole thing is about silliness and having fun and forgetting about kids and kids as problems and stuff like that and so we had um.. in fact...is everything (pointing to video camera) rolling here?

9) PP: Yes yes. Do you want me to turn it off

10) SD: No no. In fact in our copier room I have a big picture of us dedicating Manly Men Park to uh with the sign the first of November or October. it was..
and so I like people to ask me about it and I like people to actually COMPLAIN about it because they don't understand that its all about being silly and honoring .. the opposite sex and being supportive of each other.
And ..you would be surprised .. the year we had last year was unbelievable in terms of staff morale
And this is a good year too was but last year just snuck up and all of sudden "we're having a great year aren't we?" and yeah we are all getting along and we are doing things socially and having fun and being a little bit silly and they got those calendars and they went nuts I have to show you one of those too and they just went nuts and they just carried us through that tough part of the year.So we are the Manly Men.

In considering Burt's role at Swallow, I have evidence that he participated in and led much of the departmental review as well the K-12 all-faculty review. I also asked Burt about the stress level at Swallow during the MEAP test and how he responded to it. He indicated it was a very stressful time for the staff, but did not provide any details about how he helped the staff cope. When pressed for further details about the stress of that time, Burt described the tasks that the different staff members were responsible for, but did not discuss the overall emotional state of the faculty. The emotional condition was the focus of Stan's response, who said later that "MEAP is a significant stressor before and during the testing period. The level of concern is raised beyond the comfort zone of students and staff."

a) Stan Dubovski (in tie) with school teachers

a) Burt Wainwright with the science department

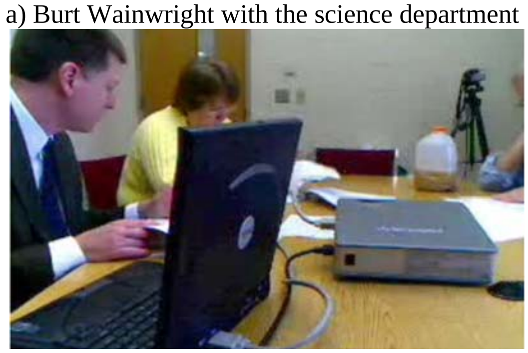


Photo 8.3 - Principals responding to the MEAP

These two glimpses of the leadership practices of these two case participants show different ways school leaders may be responding to the overall testing process. With Burt Wainwright, the response is direct and explicit. He chose to integrate the MEAP process into his job; taking ownership of the interpretation process and directing the school faculty through the steps of the review. He also did not have a curriculum specialist upon whom he could off-load this responsibility. In Avon Falls, Stan was more able to delegate the details of the MEAP interpretation, because he had a curriculum coordinator in Janey Fess who had both the position to manage the process and broad subject matter knowledge to be able to engage in instructionally relevant discussions. Stan thus was left more able to focus on creating a culture of mutual support. Burt's response focused on the kids, whereas Stan emphasized "having fun and forgetting about kids and kids as problems" (Transcript 8.10, turn 8).

Analysis of the Temporal Structuring of the Interpretation Process

This analysis builds on Lemke's (2000) timescale model by treating the test administration and score releases as points that mark intervals in the school calendar. Stan Dubovski's response to the stress of the MEAP introduced questions about how Michigan's testing process, with the administration in the fall and release of results occurring within the same school year, affected the temporal structure of the school's practices. In this analysis, I expand the scope of this question to other schools across Michigan. Looking at these as types of activities, I ask whether the relationship between state and schools is constant or whether these interaction points segment the year within the schools.

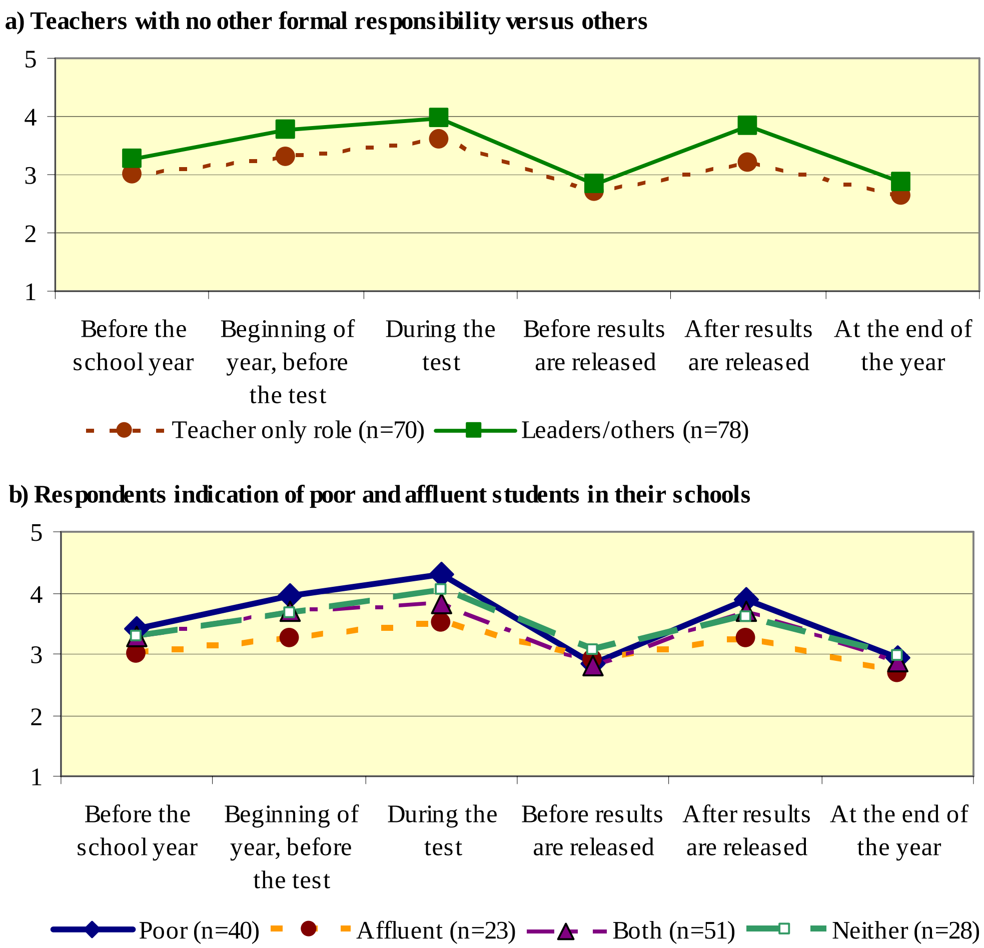


Figure 8.2 - Respondents' average rating of how much MEAP is in their minds at different periods of the year

To probe this dimension, I began by asking Burt Wainwright about the attention in the school to MEAP issues during different parts of the year. He responded that the MEAP was still a focus of his and his staff's attention between the October testing and the January release of the results, but not with the same high priority it was given after the results are released. To further understand how the MEAP affected the collective consciousness of the schools, I placed a qualitative perception question on the MSTA survey (see Chapter 6) asking the respondents to indicate the position that the MEAP had in their daily thoughts throughout different parts of the year. The lowest rating was 1 (not in their minds) and the highest was 5 (a constant focus). The responses suggest the MEAP has an impact in terms of attention throughout different parts of the year, as illustrated in Figure 8.2.

As discussed in Chapter 6, this survey is useful, in combination with other data sources, to illustrate potential patterns occurring in communities broader than those schools in the study. These responses suggest that the administration schedule, with two contact points within the school year, does affect the consciousness throughout the year for many schools. But it does do not give any indication of changes in practice beyond the likely meetings held to review the related scores.

While this general pattern did hold across all the different ways I analyzed these responses (school type, district size, and experience with the MEAP development), not all respondents indicated the same topology. Of those responding to this part of the survey, 19.7% provided the same rating for all parts of the year. These flat responses indicated no part of the year was any different in terms of the role of the MEAP in the school. Some of these flat-line responders were like Paul Bond, who said that the MEAP had no appreciable impact on them, and others were like Christy Connolly, indicating a constant focus. Still other flat-line responders were in between these two ends.

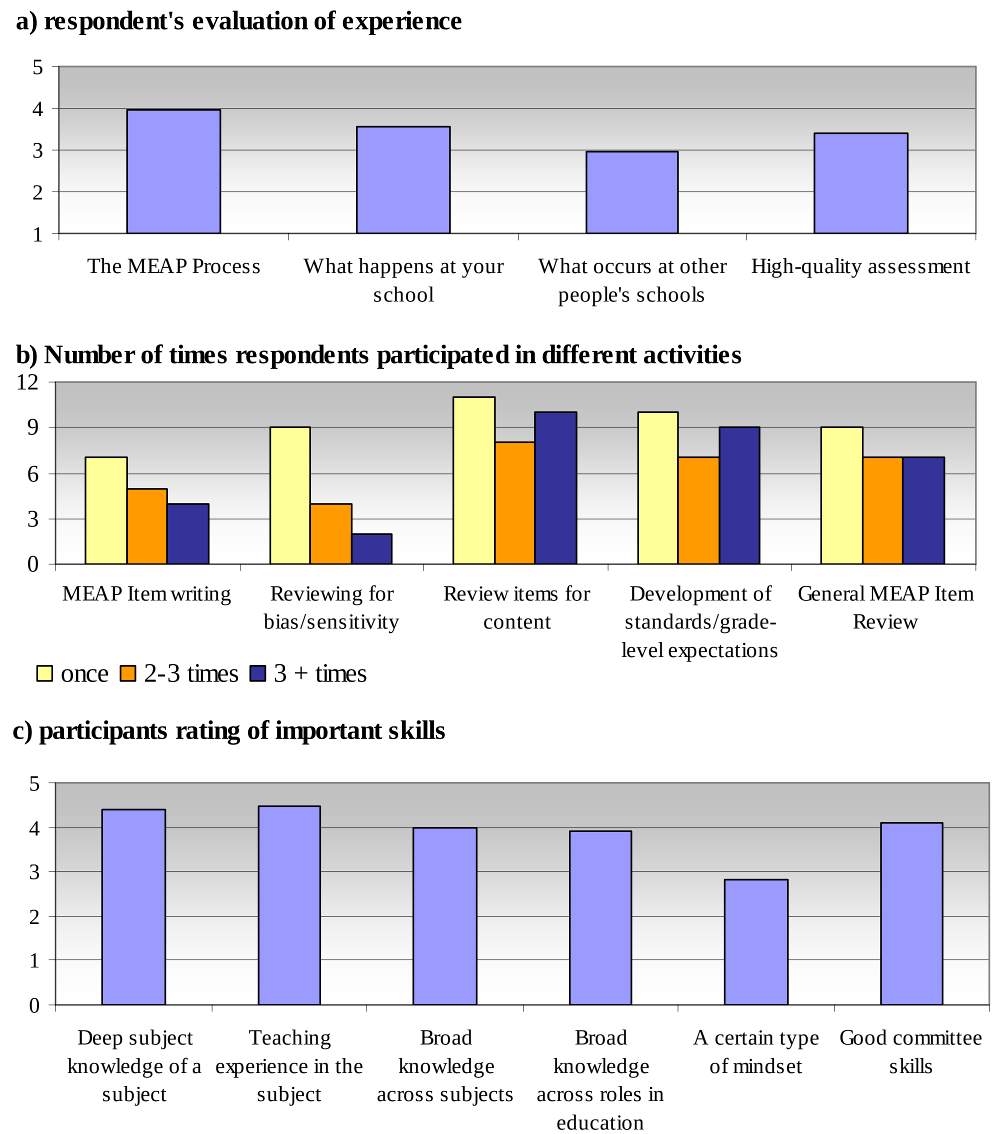


Figure 8.3 - Survey responses regarding brokering.

A Glimpse of Brokering

Individuals who participate in multiple activity systems or communities are of key interest in discussions related to organizational integration. Wenger (1998) calls them brokers. Tushman and Scanlan (1981) refer to them as boundary spanners. They complement the perspective of the artifacts that Latour (1987) referred to as immutable mobiles and Star and Griesemer (1989) referred to as boundary objects. There are several individuals whose boundary-spanning activities have been included in this study. Three of the more prominent ones are Bob Senoff, Nancy Newman, and Christy

Connolly. Christy's experience in the MEAP review process, discussed earlier in this chapter, is extremely germane to this discussion, as she is also a study teacher. As this chapter closes, I want to look at the brokering activity of those who participate in the MEAP process. While the individuals who serve in these roles are important theoretically, as they allow us to sketch connected networks, what does that role mean for the individuals? Do they, like that quote at the beginning of this chapter, see things anew because of the new places they go as they span boundaries?

Table 8.4 presents some selected comments from the MSTA survey by those who indicated they had some experience with MEAP development. While often positive, these responses show a diversity of opinions, indicating that perception of the activity is not uniform but rather varies across individuals.

Table 8.4 - Selected comments from MSTA survey on experience of working with the state.

Experience with	Comments
Item writing, bias and content reviews	Anyone who wants a say in future changes and sees teaching as a profession, not just a job.
Item writing Development GLCEs, bias and content reviews	Every teacher should participate in development/review. Teachers need to understand the process of how a MEAP test is prepared
Development GLCEs, bias and content reviews	I would like to see more teacher representation from the west side of the state and the northern lower peninsula region. Demographics should not be the final determination for input on the MEAP or with the State's GLCE's. With advance technology, distance should no longer be a hindrance.
Development GLCEs, bias and content reviews	My understanding on how good questions are written and how standardized tests are developed has deepened during the meetings. I also have a great grasp on the new MI curriculum because you are constantly referring to it and discussing the wording and the meaning of all subject areas. You also have an opportunity to share teaching ideas with teacher s from all over the State. You are treated like the professional you are who has expertise in your subject area. I always come away from the work sessions with a renewed vigor to approach my curriculum from more focused and yet more innovative angles.

Two Systems Interacting

While Chapter 7 focused on the web of local, intermediate, and state organizations that individuals frequently span, this chapter has privileged two categories: the school and the state. Further, the relationship between these two levels was analyzed in terms of the annual accountability system, a key component of the current period of systemic reform. In a sense, this chapter is at the other end of some issues raised in Chapter 5, describing the analogous assessment system from a higher ecosocial level (Lemke, 2000). And, like Chapter 5, it shows that while the explicit purpose of the assessment system is measuring student knowledge, in practice it entails other important dimensions, including school leadership and climate.

The approach Michigan uses of including practitioners in the MEAP development process yields important benefits. It enables brokers who can participate in discussions related to a nexus of educational issues, including performance standards, student cognition, teaching methods, and classroom equipment. The benefits of these practices, however, are not universally perceived. As was shown in Chapter 7, some teachers are more connected than others.

Chapter 9 A Dynamic View

"Prediction is very difficult, especially if it's about the future."
~ Niels Bohr

Introduction

This chapter takes the model that has been building over the last four chapters – a model involving a network of evidentiary components relevant to the State of Michigan – and animates it. I begin to introduce topics of change and fluidity into the study and raise questions that are important to consider in thinking about what the depiction in this dissertation might mean for broader contexts. This chapter looks at a mixture of cases and types of cases as it addresses these issues of systemic movement. The evidence across these cases illustrates the changes that had been occurring in the years leading up this study, glimpses of changes that were occurring while the study was being conducted, and some anticipation of how systemic reform and new technologies that support reform discussions are suggesting what might occur in the years directly following this study..

I begin this chapter by introducing one last case study highlight: the use of a data warehouse for Hardy Middle School. This discussion begins where Chapter 8 ended, by discussing MEAP scores. However, the way that Hardy was reviewing MEAP scores and the issues associated with what I observed at Hardy add important new dimensions to the discussion. Aligning instruction to state standards and school leadership are issues present in this case, as they were in the discussions from Swallow and Avon Falls in Chapter 8. But, the nature of Hardy's leadership and how data was being used at Hardy raise broader dimensions of equity, accountability, and responsibility.

The Hardy discussion leads to examination of trends across Michigan toward developing information infrastructures for examining practice through student performance. I call these efforts "The Warehouse Wave," because of the rapid proliferation of information technologies that is their signature feature. This wave is broad and evidenced in some of the public documents of the assessment and administrative professional community cases, along with reports from individuals, in this study.

These reports align with the literature reviewed in Chapter 2 that showed an increase in studies that considered information use in schools. Many of the sources in that review highlighted the importance of the information infrastructures that support data-oriented practices, also raised in articles reviewed in Chapter 2. In this chapter I describe a parallel trend within the State of Michigan where, across the region, intermediate school districts have been broadly initiating projects to provide student performance information to educators. I analyze this trend and characterize it as a *diffusion process* (Rogers, 1995) whereby adoption of innovations follows a pattern that includes early, middle, and late adopters. This diffusion characterization, while appropriate, is complicated by factors related to organizational and technological considerations. The chapter concludes by focusing back on practitioners. But, unlike Chapter 5, this practice discussion includes more than just teachers. It also looks at issues with leaders and teams of teachers that reflect the broadening of instructional responsibility that has so marked this era.

Case Highlight: MI-Tracker in Hardy Middle School

Just after opening the door to enter Hardy Middle School, a visitor was greeted by a sign that said, "YOU HAVE JUST ENTERED A GUM FREE BUILDING!" This sign, and its similarity to the sign that hangs over the entrance to Faith Churchill's classroom 125 miles away in Dorchester Middle School, are helpful in focusing attention on practitioners as this dissertation nears completion. As the zone of the educational system where teachers' daily work again comes into focus, this sign helps to expand the discussion beyond individual classrooms to buildings. It is at the level of the school where much of the emphasis from NCLB lies, and it is within buildings that teachers, principals, and often specialists work as a team. Just as the classroom is thought of as a locus for student learning, for assessment information use, the building can be thought of as a locus for practitioner learning and also increasingly for assessment information use.

The literature review in Chapter 2 shows (see Table 2.5) that across the early studies of data use in schools, four common factors had emerged as critical to the success of these projects: leadership, a technical infrastructure that can disseminate data in a timely manner, processes of organizational readiness (ex: data literacy), and information visualization. If we were to look for one of these factors in the way that the Hardy Middle School's review of their MEAP scores in 2006-2007 was significantly different from the way Swallow and Avon Falls reviewed theirs, it would be *visualization*. As Dianne VanderMiller prepared at Hardy for that school year, she used the same standard black and white reports from the OEAA website that were used in Swallow and Avon Falls. However, during that 2006-2007 school year, Hardy's district began to use the MI-Tracker data warehouse/student performance information system. With this system, Hardy's information visualization options increased. Note that throughout this chapter, I will use the terms "data warehouse" and "student information system" interchangeably,[46] although some differences exist in their use.

[46] From a strict technological perspective, the terms should be construed differently, as Bernie Lauer emphasized in Transcript 7.10. An even more specific difference between data warehouses and other types of information reporting systems would focus on the technologies used and the database designs employed. Data warehouses traditionally employ what is called a "star schema" design that has specific analytic properties and is optimized for complex queries.

Hardy's use of MI-Tracker places it in a small and growing group of districts that are using these types of technologies for their interpretation of the MEAP results. The MSTA survey responses showed that 15% of respondents received their MEAP information from a data warehouse; 62% indicated they used the standard reports from the state government, and 69% reported some custom reports and displays local to their school.

Comparisons between what occurred at Hardy and what I observed in Swallow and Avon Falls need to consider differences in these schools' circumstances as well as in their information technologies. At neither Swallow nor Avon Falls did the issue of those schools not making AYP come up, while at Hardy it was an important theme that surfaced in my conversations with the school and district leaders . Also, the leadership stance that Dianne and her principal, Carmen DuRonder, take is more pro-accountability than either Burt Wainwright or Stan Dubovski exhibited. Before discussing MI-Tracker and my observation of the training session for it, I will expand on the context for this discussion by briefly touching on reform efforts affecting Hardy at the time of this study.

Reform Effort in Hardy

Recalling the discussion in Chapter 7 about Dianne VanderMiller's role, Hardy was undergoing an improvement program when I was there, and Dianne held a role that was part instructional coach and part school leader. Carmen DuRonder , who had been principal at Hardy for five years, discussed her vision for the school and her changing role from building manager to instructional leader several days later. This conversation, represented by Transcript 9.1, occurred in her office with Dianne present for part of the meeting. During the MI-Tracker training session, Dianne shared her moral commitment to teach all students regardless of their circumstances. In another conversation in her office, she said, "Parents send us the best product they have. They don't keep the good ones at home." Issues of subgroup and disability also arose when Dianne discussed the poverty of the students coming into the school and the instructional challenges that they face in Transcript 9.2. This conversation occurred during the MI-Tracker training when Diane was working with Ann Pobzerniac (AP), who had been a seventh grade science teacher and was an ally of Dianne's in this school reform effort. Dianne reported her position was at odds with much of the staff. While Valerie Jones, the seventh grade science teacher I interviewed, did not comment on staff tensions in the school, Carmen and Bob Senoff both echoed Dianne's description of a school climate where accountability was a very contentious issue.

Transcript 9.1 - Carmen DuRonder (With Dianne VanderMiller) discusses vision for school, her changing role, and data being used in new ways.

1) CD: The biggest part of that is that if you look at some of the inner city schools and in upstate New York and all that good stuff, those kids don't have parent involvement.

2) PP: Right:

3) CD: Ok what's the difference between that in that? It's the school. It's the teachers. And, it's the staff.

And y'know just what makes a child at risk. And people say it's economics it's the home life.. uhuh. Its that they are not getting what they need at school that makes them at risk. We're the ones who are setting them up for failure if we are not meeting their needs.

So if we're providing kids with homework and we know they're not going to do it because they have to baby sit three kids at home and there's no parents.. we're setting them up for failure when they don't turn it in. So, what are gonna do differently so they don't do that. Its a huge y'know people think at-risk kids don't have money, and they're poor, and they have no parents.. No our kids are at risk, they come to us all kids are the same and its our job to make them not at risk for failure. We have the power to do that in school. But, it is changing the way people think.

-–—later

4) CD: And that's a tough role. I'm not I'm not here to be .. their friend. I'm here to be .. you know their leader to show them the way to go and have the respect. I'm not exactly real popular about that. But, the people who are on board the people who know you are doing good things are starting to cheer a little bit.

5) DV: Oh yes.

6) CD: And, they're starting to come out of the woodwork more. And people are starting to say "uhoh, we better do something ." So

7) PP: And it isn't so there isn't a pattern that you see in terms of [teachers who support accountability]..

8) CD: I would say the pattern would be somebody who has been in the maybe the profession for a little bit for a while. But that isn't always true because we have some really great veteran teachers who are here for kids. I would say it would be somebody that gets sucked into the negativity that can happen in this profession and the us versus them. With us, administration, versus teachers. And some young teachers get sucked into that mind frame and um we have been lacking in the manpower to look keep an eye on them and pull them out and that has been changing.

Transcript 9.2 - Dianne VanderMiller discusses Special Education and disadvantaged students.

1) DV: Like our special Ed population our free and reduced lunch is substantially statistically much higher than Northern (the district's other middle school). So that's why I was moved there [Hardy].

2) PP: Do you have much um of a conversation around Special Ed when you are also looking at all of this data.

3) DV: yes.

4) PP: Is it a special category or is it integrated is it?

5) DV: Do you mean looking at the data or servicing those kids or what?

6) PP: Looking at the data and the things you do with that data.

7) DV: That is our greatest concern.. and we (pointing to Ann) were just talking about that. If we don't make AYP that'll be why.

8) PP: Because of the Special Ed?

9) DV: Yes. And, but we (points to Ann) believe that is a systemic problem .. right here. That we need to take care of; that we need to fix. Our kids are pulled out for Special Ed in elementary, in middle school it's inclusion, they're pulled out in high school.

So these children they are pulled out for Special Ed in elementary and they are not necessarily really receiving instruction at their grade level content expectations for that grade level.

10) PP: Right

11) DV: And they are put in sixth grade, they are put into sixth grade math class where the resource teacher is a is a help and they're in the sixth grade curriculum they haven't had fifth grade, fourth grade, and in some cases third grade... and so there are these massive gaps.

It is really a systemic problem that we feel um ..

12) PP: And you see it as an issue with here ... as something that can be changed? this is..

13) DV: I think, well, I do. (looking at Ann) I'm not sure if everyone would agree with that.

14) AP: It can be changed, it's just is going to have to be the district that changes.

15) DV: But I'll say, Joe Crang, who is our special ed supervisor? I don't know the word. I'm not sure. He and I have had multiple conversations about this and he he is quite passionate about servicing those children better.

And I'm not sure that um I think maybe this has been .. not just here but .. think it seemed people felt we could do the best we could and we did what we could for those children until they started counting for AYP and suddenly they started mattering a whole lot more. And that's just but that's just my own personal opinion.

The reformist nature of Carmen and Dianne's roles emerge in these discussions. Carmen's description of the school's overall responsibility to teach all kids (Transcript 9.1, turn 3) parallels Dianne's position that the school needs to work with those kids that they have. Rather than "at-risk" being constructed as an attribute of a child, Carmen treats it as a process failure within the school (turn 3). The low level of homework shown in Valeries's practice in Chapter 5 (see Figures 5.7, 5.8) can now be interpreted within this larger context of homework being viewed as accentuating inequity at Hardy. In fact, Valerie had indicated that most homework she assigns is work that was not completed in class, rather than additional work.

MI-Tracker Training

MI-Tracker is the product of a small firm called Successline, Inc. based in southeastern Virginia. At the time of this study, Successline's software products were targeted to two states, Virginia and Michigan. As the selected image from the Successline website shows (Document 9.1.a), the two versions of their product were advertised to be compliant with both state requirements and NCLB. There was no mention of broader support for other measurement systems beyond state tests in the company literature. Mike Wolson, co-founder of Successline, told me he had been a captain in the United States Navy working with management information in the Seventh Fleet before he met his wife Elizabeth and founded the company. Elizabeth had worked with middle school science before earning her doctorate in urban studies from a small Virginia university. The company website named them both as the only management contacts and listed services and books that they provided in addition to the software. Mike also told me that the three offerings from the company – software, services, and books – worked together as a complete package.

Successline was shown earlier in Figures 7.1 and 7.2 as linking various schools in this study. Stan Dubovski at Avon Falls gave me one of Successline's books, titled *Designing High-Quality Paper-and-Pencil Tests* (Successline, 2002). The book describes basic ways of developing multiple-choice items, and Stan told me that he got the book as part of an ISD-sponsored workshop where Elizabeth was the facilitator. Successline is the company that ran the workshop discussed by Pete Darmond in Transcript 6.1 that included representatives from both his McReady School District and the other school systems.

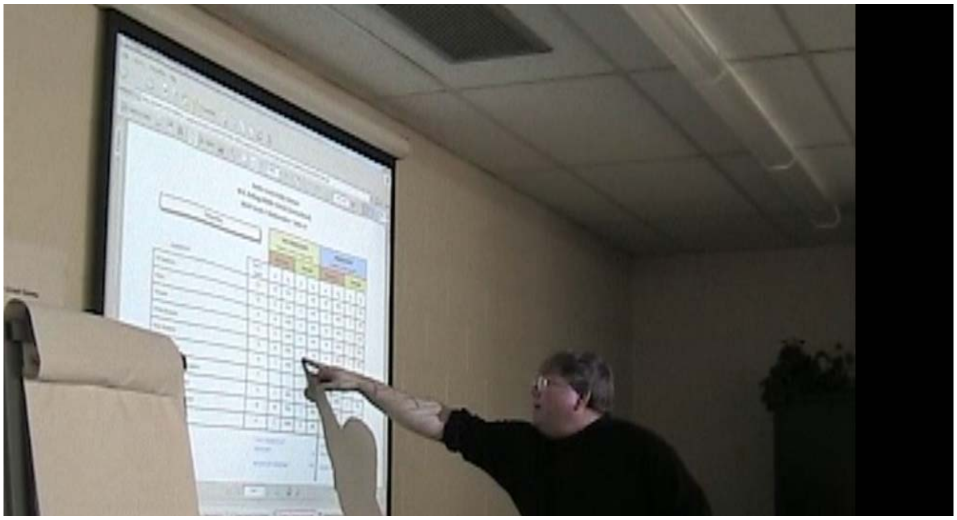


Photo 9.1 - Mike Wolson leading MI-Tracker training

I observed a MI-Tracker training session that brought together over a dozen teachers in Hardy’s district. The attendees came from different buildings and taught different subjects. In this session, shown in Photo 9.1, Mike Wolson (MW) led the group through exercises where the teachers used the MI-Tracker reports to diagnose potential instructional problems illustrated in their MEAP results. In these exercises, the discussions included a range of subjects not covered in the reviews in Swallow and Avon Falls. In addition to topical alignment and science teaching methods, these discussions highlighted subgroups, an important component of NCLB. Subgroups under NCLB include students from disadvantaged socioeconomic categories (ex: African American) and those with disabilities, as Transcript 9.3 illustrates.

Transcript 9.3 - Mike Wolson presents a MI-Tracker report to diagnose instructional breakdowns.

What I've got is grade seven and grade eight math and students with disability. Low performing sub-group right? I mean I don't even have to look to see what the numbers are. I know that students with disabilities is below the AYP limit as far as that goes.

So here's what I've done is I've gone down into the grid and if I total up the number if students tested in this building, total number of students, students with disability, it comes up to 21 students. It's below the state for a legitimate subgroup. But, guess what? They've probably got some MI-ACCESS kids in there. That group could grow in the next two or three years.

And what would happen is that they’re now in trouble. I'm not saying they’re getting away, but the accountability piece is not at the, held at the building level. It may be held at the higher level, at the district level. But because they

Transcript 9.3 - Mike Wolson presents a MI-Tracker report to diagnose instructional breakdowns.
are below 30 now there is no consequence.
If they grow, they're in trouble if they don't fix this problem.
So that's the third example of saying look if I don't deal with this subgroup right now down the road it is going to be a problem for me - so I need to take a look at it.

In Transcript 9.3, Wolson is describing how summaries of assessment data that are broken down by subgroup can be used to identify a potential problem with a group of students. While the standard MEAP reports from OEAA also include similar subgroup breakouts, those reports are in paper form and may make the analysis cumbersome, because the information needed for this analysis is spread over several reports. In contrast, MI-Tracker reports use color and contain various summary levels and disaggregated results at more analytic levels, leading to a possibly more robust set of data conversations.

In his presentation of the disability subgroup, Mike highlights two important dimensions. The first is the group's composition: perhaps there are some kids in the group with cognitive disabilities who would use the MI-ACCESS assessment system from the state government. MI-ACCESS (no relation to MI-Tracker) is developed and administered by an OEAA group that is parallel to the regular MEAP group. The MI-ACCESS team includes Patty Hollander and Vance Dorn, who were informants in the study discussed in Chapter 7. The second dimension is the trend over time of this group. Mike points out that with just a few more students, this group could become a future problem down the road for Hardy. Neither of the MEAP reviews discussed in Chapter 8 had much focus on subgroups or much focus beyond current results, whereas both were discussed several times at Hardy.

Document set 9.1 provides some overviews of a few of the reports available from the MI-Tracker system, showing reports that group students by strand (b) that display the topology of student achievement for a particular constructed response item (c), and that compare students in the school to district and state by benchmark (GLCE) within strand (d).

a) Product links on web site

Software

MI and SOL Tracker is state

and No Child Left Behind Act Compliant

View Information on MI Tracker for Michigan Educators

View Information on SOL Tracker for Virginia Educators

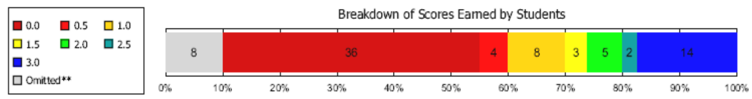
b) Sample strand report

STRANDS	Total # Tested	NOT PROFICIENT Students in Level 3 or 4				PROFICIENT Students in Level 1 or 2			
		Weakness		Strength		Weakness		Strength	
		#	%	#	%	#	%	#	%
Construct New Scientific Knowledge	80	52	98%	1	2%	19	70%	8	30%
Reflecting on Scientific Knowledge (R)	80	31	58%	22	42%	4	15%	23	85%

c) Sample display of constructed response item results

Earth Science

Benchmark*	Item Descriptor	Maximum Possible Points	# Students Answering	# Students Earning 80% of Possible Points	% Students Earning 80% of Possible Points
SC.5.48	Provide examples how specific earth materials are used by mankind.	3.0	80	16	20%



d) Section of sample report that compares item performance to others in district and state grouped by strand

Comparison of Item Analysis Results between School, District and State

Pleasantville Public Schools
Pleasantville Elementary School #1
MEAP Grade 5 Science • Winter 2003-04 • Form B

Section	Item Descriptor	# Students in School	% Correct in School	% Correct in District	% Correct in State
Construct New Scientific Knowledge					
SC.5.01	Identify the simple machine used in exercise equipment.	81	69%	62%	82%
SC.5.02	Develop a solution to identify whether an object is made of iron.	81	63%	81%	88%
SC.5.10	Selection of an instrument to measure temperature.	81	96%	97%	99%
SC.5.11	Selecting the appropriate unit of measurement for a task	81	58%	63%	75%
SC.5.12	Note measurement instrument with least application for investigation.	81	19%	19%	26%
SC.5.13	Using results to recognize a conclusion from the investigation.	81	60%	68%	77%

Document set 9.1 - Examples of Successline, Inc. advertisement and reports

Another issue that Wolson showed the group how to analyze was the alignment of instruction. He used MI-Tracker reports to identify problems with individual teachers or groups of teachers (ex: a school), as Transcript 9.4 illustrates. Unlike the item-based reviews discussed in Chapter 8, Mike used the strands and GLCEs to structure the report and the conversation. In this review, the role of individual assessment items also changed. While Chapter 8 described processes that were largely structured by the MEAP items, in Hardy the role of items decreased, while the role of superordinate categories of standard/GLCE and strand increased. Rather than looking at which questions were answered incorrectly independent of standard or strand, MI-Tracker provided reports (see Documents 9.1.b/d) that summarized at the higher levels. At the end of this transcript, Mike points out “the AH-HA piece” is when the reports helped to identify the broken (unsuccessfully taught or not taught) GLCE.

Transcript 9.4 - Mike Wolson presents using reports to identify “broken” GLCEs

What we're trying to is get you to work on is the alignment.

That (pointing to part of report on screen) is telling me, those numbers are telling me that I have an alignment issue. I'm either not teaching the state curriculum or what I am teaching, I am not teaching the right stuff I'm teaching the wrong stuff really well, or I may not be teaching it at the right time.

Ok that's what the whole root cause business is and when we get down to the back end of the process here it's going to be clear to you. You will have .. you'll know why because you've got the reports and you'll know what GLCEs are broken and all that.

--- Later ---

Here's the hardest part of this whole process because this is what gives you the root cause of the problem.

I'm looking at it by GLCE so, I've got the GLCE now prioritized that I want to work in.

Don't go to the question. Go to the GLCE and say What is the concepts, are the concepts and the vocabulary required for mastery of that GLCE.

In other words you need to unpack the GLCE. If you guys have not unpacked the GLCE then it's not going to happen this morning, obviously. But that's what the whole point is.

Is that if I don't know what the vocabulary is and I don't know what the concepts are I certainly can't point to them in the curriculum guide or my pacing guide. I can't find any of that stuff and that's why the GLCE is broken and that's the AH-HA piece.

Not only does Mike emphasize the GLCE can be broken, he also discusses the process of interpreting the standard that involves what he calls “unpacking” and others have referred to as “clarifying” the standard (Harris, et al., 2003). Considering science standards specifically, it is also important to note, as was discussed in other schools where MEAP results were reviewed, that Michigan had not been testing all of the GLCEs every year. This may present additional alignment challenges for science teachers as they attempt to guess what GLCEs are most likely to be tested.

Information seemed to be integral to the work that the leadership team at Hardy performed. When asked about how data helped in working with teachers, Dianne responded that it provided “massive leverage.” Carmen shared (in Transcript 9.5), as did many others in the study who work at the intermediate and state levels, including Bernie Lauer, Nancy Newman, Joshua Martinique, and Pete Darmond, that the two to three years before this study were a time when the concept of using information for instructional purposes gained prominence. As Martinique, who worked in one of the largest ISDs before joining the state government, remarked, “We always had MEAP results around and were analyzing it, but in that time [two-three years] we started to see a change.” Carmen describes her experience in Transcript 9.5.

Transcript 9.5 - Carmen DuRonder (With Dianne VanderMiller) discusses vision for school, changing role, and new ways data used	
1)	PP: Thinking back on your role as a principal, five years ago what was the relationship what was the role of data to the the type of work you do here or at the high school.
2)	CD: I think back five years ago it wasn't .. y'know it was there and talked about, but it wasn't really addressed. And my role was mostly dealing with student discipline my role was dealing with master scheduling so I'm always looking at number data and credits and things like that. But to actually be brought down about test data and to be held accountable for it. That's really something that's really changed with the AYP in the last few years.

I now broaden the study’s aperture from a particular school case to districts across the state to consider how what was occurring at Hardy might fit within a larger process related to the instructional use of information occurring within Michigan during this time. MI-Tracker belongs to a category of information systems that were prevalent in Michigan during this period, but scarce just a few years earlier. Chapter 2 showed a recent interest in studying these types of processes. But the literature reviewed did not focus on indigenous processes across a region, like a state or county. It was not clear how indicative that literature was of trends in practice or whether it was largely driven by academic and research motivations. In this short analysis, I will draw primarily from the school administration and testing professional association cases as I look for evidence of a trend in the rapid proliferation of software systems that help collect, manage, and distribute assessment information across districts and intermediate districts in Michigan.

Analysis: The Warehouse Wave

In 2005, the technology committee of MAISA conducted a survey of its membership to ask about their plans for adopting data warehouse technology. Later, in 2007, MAISA participated in another survey, managed by MBSA, that covered the same general questions: Which ISDs were undertaking data warehouse efforts? And what were the statuses of those efforts? These two surveys show important changes in the landscape of technology infrastructure across the state. In 2005, 38 (of 57) ISDs responded to the survey and 25 indicated they had a data warehouse project underway. Half of those projects in 2005 were in the initial planning stages, while 13 were in the early roll-out stages. None had been implemented for more than three years. By early 2007, the number of respondents increased to 40 and the number of reported projects also increased, to 30. The number in various stages of their life cycle from conception to implementation also advanced as well:

> 23 of the 40 indicating their various stages of data warehouse implementation identified the number of years the project has been in place: 11 are in the roll out phase, 4 indicated 1 year, 6 indicated 2 years, 1 indicated 3 years, and 1 indicated more than 3 years. (MAISA, 2007)

While these two surveys are an important category of evidence to support a trend toward widespread advances in district/ISD technology in Michigan during this time, across them are important differences. The types of questions differed, and the published reports differed in some ways also. Furthermore, five of the ISDs present in the first survey did not respond to the second. The ISDs responsible for Swallow, Avon Falls, and Hardy schools responded to both surveys as did Maple Grove, where Bernie Lauer works, and McReady, where Pete Darmond works. Washtunesa ISD, where Nancy Newman worked and is responsible for Dorchester and Crimson schools, responded to neither. This is consistent with there being no mention of data warehouses in any of the interviews associated with this ISD or its districts. Tichiochi, the ISD where Karen Minor and Don Pulte work, responded to the 2007 survey only, even though there is evidence that the ISD had been pursuing data warehousing for several years and Don Pulte was on the technology committee.

It would follow that these efforts might also appear at some point in the professional community for assessment, evaluation, and testing. To explore this possibility, I reviewed the conference programs from five Michigan School Testing Conferences (MTSC) between 2003 and 2007. I performed an informal content analysis on these programs where I attributed each session to one of eight categories, as shown in Table 9.1. The categories used in Table 9.1 are not pure, and there were cases where a given presentation could have been attributed to more than one category. However, since my goal in this analysis was to identify some evidence of impact from these data warehouse projects on the assessment community, the overlaps were deemed of slight consequence. During this period, the MSTC began to have presentations related to this topic as well as some related to district assessments/alternative assessments, a topic that was raised as a factor in several of the studies reviewed in Chapter 2.

Table 9.1 - Categorized topics of Michigan Science Testing Conference sessions

	2003	**2004**	**2005**	**2006**	**2007**
MDE: MEAP, MME, updates	14	8	11	11	8
NCLB & national issues	5	4	4	5	6
Statistics, psychometrics, and measurement techniques	2	5	4	2	4
Data warehouse and infrastructure	**0**	**2**	**0**	**1**	**2**
District, interim, and alternative assessments	2	3	5	6	2
Disability issues (including MI-ACCESS)	1	1	1	4	3
Subject area specific (Literacy & Math)	6	5	4	4	5
Vendor and other presentations	10	9	7	3	3
	40	37	36	36	33

Over these five years, five sessions had descriptions that showed a relationship to the data warehouse survey. One of two presentations in 2004 was actually by the state Office of Educational Indicators, discussing data available at the state level, and the other one in 2004 was about general data warehouse design topics. Titled "Whetting Your Appetite for Data," it was a description of how to design a system, rather than a report of an actual project. This is consistent with the Data Warehousing Surveys that would have shown only a few systems in Michigan that had been in operation long enough to report results. The first real project presentation was in 2006. It was from the Scorpio Bay School District, which was implementing a common assessments system. Scorpio Bay is a district within Bernie Lauer's Maple ISD, and he was discussed as a key supporter of/mentor to that project. This presentation was a status report from the early stages of their effort. In 2007, there were two sessions with four presentations that actually related to usable information systems. One was a session on the Scorpio Bay project where the presenters discussed the system as it was then working in the high school. The other had three individual presentations from three different ISDs discussing their experience with operational systems. One of these operational presentations was by Bernie Lauer and was represented earlier in Transcript 7.10. The other two were presented by associates of his, who I found had presented with him at MSTC and MERA before.

It is important to note that the MSTC is not a peer-reviewed forum. Rather, it is managed by a committee that includes representatives from the University of Michigan, MERA, and MDE, along with others in the testing community. A planning committee meets to decide which types of sessions to ask for and that group had included a small number of individuals until the recent retirement of a long-time organizer (Roop, personal communication, March 26, 2007). The planning for the 2008 conference used a new organizer, but many of the same individuals and organizations participated as they had in previous years.

Connecting to Local Case Evidence

Even as the focus of this discussion has been general and looking at processes across the state, it relates to particular localities within this study as well. Fifteen of the MSTC presentations in 2003-2007 listed as speakers individual who contributed to this study as case and non-case participants. Bernie Lauer, a member of the organizing committee, presented five times. Joshua Martinique presented or was part of three sessions. Nancy Newman presented once in a session that was promoting alternative ways to assess student learning beyond the MEAP test. None of the presentations were made by science educators, with the exception of Roscoe Ellis from MDE, who presented along with other subject area representatives in a general MDE/OEAA update session.

When I asked Burt Wainwright at Swallow Middle school about systems that were alternative to the MEAP and grades, he responded (see Transcript 9.6) that a county-wide assessment effort that will help them to have interim assessments and more proactively address student learning was underway at the ISD. We discussed this issue again without the recorder while walking to his office, and he further shared that his ISD planned to implement a data warehouse solution. In 2007, some comments by ISD respondents were included in the released report. The comment for the ISD that Swallow fell within was “Partially Implemented, solution chosen and initial data being used by some of the school districts.” This characterization is fully consistent with Burt’s local account at the very start of the study. Bob Enspania at Avon Falls had also reported that his ISD was preparing to work with Cummins' ISD on a similar data warehousing project.

Transcript 9.6 - Burt Wainwright discusses his expectation of a data management system

1) PP: Last time I was here you showed me basically the MEAP results
2) BW: Umhumm
3) PP: Are there other science related information sources you have?
4) BW: Not that we do on a district uh basis for high school
5) PP: um Grades are just grades?
6) BW: Right, uhh although we are attempting to move toward some county-wide assessment types of thing so that we will have .. um not that we have to do that way it's not mandated or anything but it would be a way to have things more benchmarked to state expectations and then we see more along the way .. in a more **formative way ..** where we're at instead of waiting for the end or waiting for a MEAP um type test so um we're trying to move in that direction. We're not there yet but that's the direction we're headed.
7) PP: Some people see them as different, grades and assessments as two different things um
8) BW: And I think **that's right**, however, that's what we have at the time at this time for assessment.

The ISD survey responses and the annual testing conference presentations suggest that broad organizational commitment to programs and infrastructures for using student performance and assessment information was occurring during the time of this study. While the evidence is coarse, sometimes coming from secondary sources, it is consistent with a wave of interest and projects in information systems focused on student performance across the state that aligns with the literature reviewed in Chapter 2.

Explanation of Change: A Diffusion Model by Organization & Topic?

Recognizing the broad movement towards the use of information in organizations between the school and the state, is there a way to conceptualize this activity that helps support inferences about other, future contexts?

Rogers' (1995) *Diffusion of Innovations* suggests a starting point. In this popular book[47] (originally published in 1962), Rogers describes various approaches to studying

[47] A Google Scholar search on this title showed almost 14,000 citations for the 2003 edition.

how products and ideas diffuse through a population. He describes some of the competing and alternative diffusion frameworks and highlights some of the challenges to various diffusion theories. Two important parts of his discussion are his elaboration on Bass' (1969) new product growth model and a focus on social networks. In the predictive Bass model, which has a somewhat distinct research community, an innovation begins slowly with a set of early adopters and then rapidly accelerates through the majority of the population. The growth can be shown on an s-curve with a slow beginning and then a rapid acceleration, as as illustrated in Figure 9.1. Rogers' focus on social networks in the take-off stage emphasizes the roles of highly networked individuals he called opinion leaders. He also discusses both individual consumer diffusion processes and the complex processes that occur within organizations.

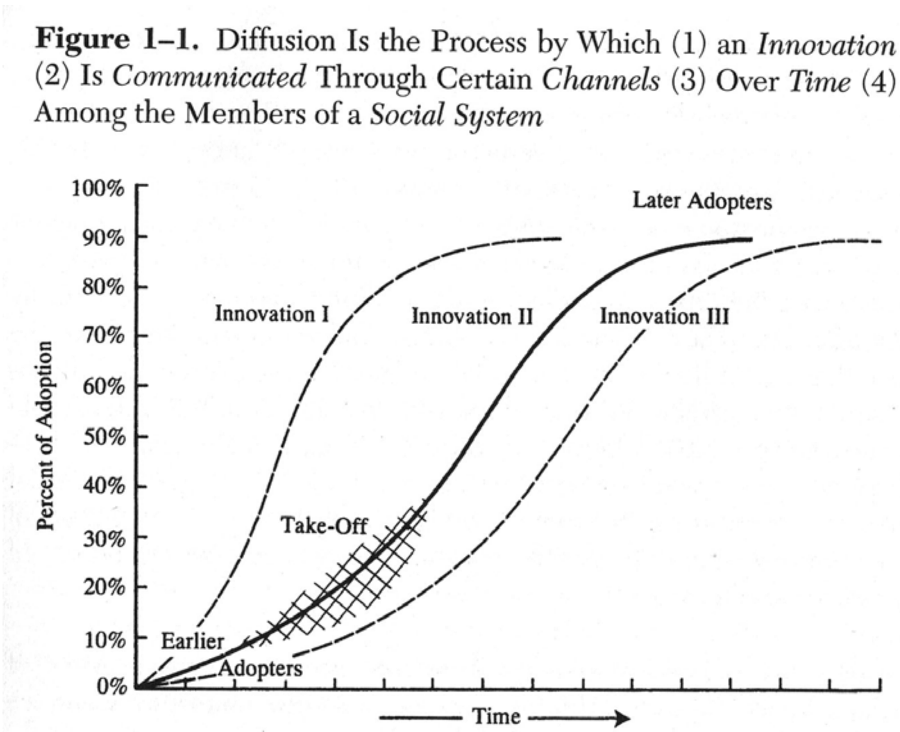


Figure 9.1 - Figure from Rogers (1995, p.11) showing three innovation adoption curves

Support for a diffusion model based on Rogers to explain the emergence of data warehouse technology reported above comes from several sources. First, the ISD surveys and MSTC conferences show that at the time of this study, there had been the widespread adoption in increasing numbers, consistent with the take-off period shown in Figure 9.1. The 2005 Data Warehousing survey indicated that there were a few projects that had been in operation for a few years and that the majority of the projects at that time were just beginning. By 2007, there was a shift towards many projects in more advanced stages, and the MSTC began to see presentations of actual project results. Second, two of the mechanisms discussed by Rogers as ways that diffusion takes off were present in Michigan during this time. The first was social networks (shown as influential by cross-hatching in Figure 9.1), including opinion leaders like Bernie Lauer and close associates of his, either within his ISD jurisdiction or t who had been frequently associated with him in MERA and MSTC. The second mechanism was educational policy ("intervention" for Rogers) in NCLB that provided a general motivation for sharing and using student information in ways not nearly as strong before enactment of the types of consequences for performance that NCLB introduced into U.S. education. This is a network mechanism that was discussed in Chapter 7. A third support for Rogers' model comes from individuals across the study, including Carmen DuRonder (Transcript 9.5) and Burt Wainwright (Transcript 9.6), who describe the rapid acceleration of the practices of using data for instructional purposes for the years leading up to this shift.

Transcript 9.7 - Don Pulte describes the broad movement towards data warehouse efforts.

1) PP: If we say three years from now, or five years from now, it's going to be a different ballgame.
2) DP: I would put one asterisk. For those that can afford to do this it's going to be a different ballgame.
3) PP: OK.
4) DP: And for those that haven't done it yet their gonna feel great pressure to do something.

   And I I already see that momentum occurring.

   As we make more and more presentations and I fielded more and more phone calls about what we are doing...And Cummings as well, they're they have a project that is about a year ahead of us as well as far their level of deployment with their schools.

   Ours is a little different in that fact that everybody is on board with the same framework. There isn't a local [district] that is out.

While the broad outlines of a widespread set of changes appeared in this study, there was also evidence of particular variations as different localities pursued different approaches. Don Pulte, the Associate Superintendent at Tichochi ISD who managed the

2007 Data Warehousing Survey and who is a leader in the DATA4SS that Karen Minor described, presented this view in Transcript 9.7.

Don, who serves in multiple roles that include being a senior ISD leader, a leader in the DATA4SS collaborative state grant discussed in Chapter 7, and chair of the technology committee for MAISA, responded to my question with both the specific nature of his ISD's implementation, which was to cover all eight of the districts in the ISD, and a sense of trends across the state (turn 4). He also highlights the fact that there are other constraining issues, such as cost and resources. The question I asked in (turn 1) was actually a follow-up to a question asking what types of districts were leading and what types were following in the general adoption of data technologies. His answer about resources (turn 2) can be seen as also an explanation for those districts that were not as advanced. Don also indicated that while Cummings has an effort that has been in schools longer than Tichiochi's, in the case of his ISD, they took a longer approach in order to include all of their local school districts. Reflecting back on the description Pete Darmond gave regarding the workshops and support his ISD gave to the districts and the lack of evidence that any changes had reached into classrooms (Transcript 6.1), Don was pointing to a complicated relationship between ISD programs and changes in practice. Certainly in the teacher and school cases in this study, there was scant evidence of classroom instruction being driven by test information, suggesting that a broad wave of data warehouse systems does not equate to similar changes in practice.

Challenges to the Diffusion Model

The diffusion model supports a set of inferences about Michigan during this time and their relationship to science education. A straightforward progress narrative that explains the difficulty in finding alternative science assessment practices is available here. By simply adding the subject dimension onto the diffusion process and considering diffusion to happen first with math and literacy, as required by NCLB, the lack of science assessment alternatives is readily explained. The review of the MSTC conference programs also reinforced the distance that science education has had from these processes, as none of the 24 subject area presentations focused on science. Twenty of the presentations were about reading or English language arts and four were about mathematics. Reinforcing this explanation, participants from all levels of the study, including Don Pulte, Stan Dubovski, Betsy Dearing, and Christy Connolly, all expressed the same sentiment: that science suffered as a result of the narrowing of efforts down to those two core areas where NCLB had applied instrumental pressure. According to this narrative explanation, should federal legislation require science to be included in NCLB, then diffusion of the science components would likely follow.

This explanation, however, contains three significant shortcuts or challenges that I will address directly. Two of these shortcuts are independent of the subject domain. They are related to the landscapes of organizational and information infrastructure in the state at the time of this study and possible trends. I will discuss both of these first two shortcuts/challenges below. A third shortcut relates to the important role that practitioners play in diffusion. Since this practice dimension is more focused on science education than the first two, I will discuss it in the last section of this chapter.

Michigan's Educational Organization Landscape

The milieu of educational organizations in Michigan that was illuminated by the evidence in Chapter 7 presents two additional complications to the diffusion explanation. The first has to do with Michigan's irregular structure, with many different organizations exhibiting variable types of relationships between schools and the state government. The second stems from pressures on these intermediate structures to possibly consolidate and realign some of their service relationships, which would have implications for data services.

Rogers (1995) articulated two primary varieties of his diffusion model. The first uses an end consumer perspective and the second is internal to organizations. In the end consumer perspective, which can include either individual people or organizations as consumers, the analytic view of the adoption process treats the selection of an innovation as an end point. The consumer perspective does not theorize what occurs subsequent to adoption. In the organizational perspective, the intra-organizational process of adoption is considered. The intra-organizational model looks at the roles that individuals within an organization play (ex: as champions), the ways that an organization can adapt and modify innovations, and the processes for developing organizational routines around the innovation.

Both of Rogers' perspectives apply to Michigan's organizational landscape during the time of this study. There is the terminal adoption of accountability methods and tools by organizations and individuals within those organizations as organizational routines (ex: annual MEAP review processes). And there are individual professionals' acceptance of systemic alignment responsibilities (ex: teachers stance on accountability or principal as instructional leader to state standards) that is illustrated in Table 7.1. There are also the complex and under-studied processes of inter-organizational diffusion. When an ISD adopts an innovation, what is the relationship between the innovation it chooses, its implementation methods, its dissemination strategy, and adoption downstream by districts, schools and teachers?

However, within Michigan's educational enterprise, we see that the network is neither purely individual nor purely organizational, but often a combination of the two. Michigan's composite organizational structure presents a more complex terrain on which to consider organizational diffusion of innovations. For example, if an ISD decides to adopt a testing technology, the districts may or may not also use this innovation. Don Pulte's comments, shown in Transcript 9.7, illustrate this point. When he indicates that his ISD's project took longer than the project in Cummins ISD and that they expected all the locals to be on board, he is giving us a glimpse of the interplay between the approach a particular ISD adopts with this technology and possible resistance to its subsequent use. Further reinforcing Don's point, since Cummins is the ISD for Avon Falls, this study confirms that not all of Cummins' districts were included in their project, since Avon Falls was only beginning to access the Successline material. Furthermore, the ISD for Hardy had at the time of this study not yet selected a product, and Hardy chose to pursue its implementation of MI-Tracker at the district level. The multiple and irregular layers within Michigan's educational system were made even more complicated by the Math/Science Center Network, which introduces inter-organizational adoption processes specific to science.

This multilayered and irregular structure is not fixed, however. The relationships between ISDs and each other, between ISDs and districts, and between these organizations and Math/Science Centers are evolving. One indication of potential movement in this area is Public Act 63 (PA 63), signed in September, 2007. PA 63 requires districts and ISDs in Michigan to conduct studies that identify opportunities to share resources and programs in the interests of reducing overall costs. As the journalist Weisenbach wrote in October, 2007:

> The goal of PA 63 is to have school officials from local districts and ISDs work together to save money and avoid duplication of services. Some of the non-instructional services schools are required to address in their reports include student transportation, human resources … legal services, food services and *technology support services*, among others.
> (emphasis added, Weisenbach, 2007, p. 1.)

While this requirement does not directly specify consolidation of educational organizations, in her article Weisenbach shares the views of several state educators who raised that possibility and discussed the consequences of consolidation. A contemporaneous effort to address the structural variation in Michigan during this time period was enacted a few weeks later as Public Act 101 (PA101). PA 101 required ISDs to adopt common calendars for all of the districts within their jurisdiction by the end of the 2008-2009 school year. A rationale for this law was the cost savings associated with common instructional programs, including the sharing of bus services. Similar to the Data Warehousing surveys, in 2006 MAISA conducted a survey of ISDs on the topic of common calendars. This survey had 38 responses, with 20 indicating some form of common calendar effort at the ISD.

These contemporary efforts, combined with the description of changes in the Math/Science Center Network that Pete Darmond introduced to this study (Transcript 6.1) and for which there is other evidence not presented in the interests of space, highlight the contingent nature of Michigan's organizational landscape generally and other contingencies associated with science education, including new science standards (GLCEs) scheduled to take effect in 2009 and new science curricular structures called

Learning Progressions that are being developed in national forums (Wilson and Bertenthal, 2005; Piety, 2007.)

Michigan's Student Performance Information Infrastructural Landscape

Another shortcut in the Rogers model that is related to the organizational diversity and evolution in Michigan is the landscape of software systems in use across the state during this period. While the ISD Data Warehousing survey indicated broad adoption of student performance systems, it also showed diversity in approach. The ISDs varied in terms of both the software products and the programmatic structures they created for them. As Table 9.2 shows, MI-Tracker is only one of many products that ISDs chose. In fact, while it was reported to be selected by the most respondents in the 2005 survey, it was in third place a short time later, in 2007. In 2005, one of the second most popular choices was the QSP product reviewed in Chen, Heritage, & Lee (2005). The other was a product from the educational publishing conglomerate Pearson. Between 2005 and 2007, Pearson's product moved from a tie for second with three selections to a tie for the top spot with five. It shared the first place position in 2007 with a product called Data Director that had been recently purchased by another publishing concern, Harcourt.[48] Both Pearson and Harcourt also had contractual relationships with the MDE during this time to support activities related to the MEAP test.

According to Rogers' network diffusion model, some individuals in a network are opinion leaders who are able to influence the decisions of others in the network. "This informal leadership is not a function of a person's formal position or status in the system. Opinion leadership is earned by the individual's technical competence, social accessibility, and conformity to the system's norms." (Rogers, 1995, p. 27.) The Pearson product suite was chosen by Bernie Lauer, who in his 2007 MSTC presentation (spotlighted in Transcript 7.10) included frequent advertisements for the Pearson product. Of the four presentations in two sessions at the 2007 MSTC conference, two (Bernie's and the Scorpio Bay project) featured Pearson's Inform.

[48] The acquisition was announced in late 2007.

Table 9.2 - Data warehouse products selected by respondents

August, 2005 Survey	March, 2007 Survey
9 chose MI-Tracker	5 chose Achieve! Data Solutions: Data Director
3 chose Inform by Pearson	5 chose Pearson Benchmark and Inform
3 chose Quality School Portfolio (QSP)	4 chose MI Tracker
2 chose Achieve Data Solutions m	2 chose Excelsior DDA (District Data Analyzer)
1 chose Skyward	1 chose TetraData Decision Suite
1 chose APS (Accountability Profile for Students)	1 chose School City and Schools Open
1 chose School City STARS	1 chose Nusoft Data Mart
1 chose CELT Corp. Group 1 Sagent Data Bus Tool	1 chose Datawise, Inc[49]
1 chose Edmin	1 chose Custom using OtisEducation Portal with Sagent Data Flow Services
1 chose Home Grown integrated to Class Server	

While the opinion leadership of Bernie Lauer provides support for Rogers' model as applied to Michigan and Bernie's agency can be seen as important for which products are ultimately chosen, other important questions about how to apply the model to Michigan during this time period remain.

The wide number of products used by the Michigan ISDs highlights the importance of this state network case study to contribute to an understanding of how inter-organizational diffusion happens in a climate of high innovation and structural fluidity. This is a climate where, like many high technology product markets, early periods may be marked by many competitors and rapidly evolving features that lead to stages of consolidation and some product stability.

Several years after Bass' (1969) seminal piece a review study recommended a new research agenda for diffusion studies. These researchers, Mahajan, Muller, and Bass, traced the evolution of the field in 1990 across three previous decades. They suggested (Mahajan, Muller, and Bass, 1990) ten new horizons for research. I will touch on just five of these research directions that are germane to the diffusion of student performance information systems in Michigan during the period of this study.

The first of these factors is price. It is clear that this study was conducted in an era of reduced financial resources for educational organizations. Less costly solutions will likely be more attractive. The second is product design. MI-Tracker was a product

[49] This product is not the same as the Data Wise publication (Boudett, City, and Murnane, 2005) reviewed in Chapter 2.

designed for Michigan standards and for the MEAP. It provided a ready solution for focusing on the results of the annual test. However, it did not have the immediate support for additional measurement systems as other products. like Pearson's Inform, that are marketed across states and require flexibility to support a wider range of data schemes. The tension between ready solutions to requirements like AYP and more comprehensive efforts emerged during this time[50].

Transcript 9.8 - Don Pulte discusses complex of issues related to data warehousing efforts

1) PP: Thinking back over the last ten years can you think of about when this really became like a really hot topic?
2) DP: Data warehousing, specifically and trying to use it to make decisions, probably in the last two years to be honest with you. It is something that we have here, this is my seventh or eighth year here of trying to get something off the ground here, and we actually three and a half years ago we were just too far ahead of where our schools were at and only one district took off. We were just too far ahead of where our schools were.
3) PP: Well you know it is interesting that say two years because almost across the board it is either two or three years.
4) DP: And if you look at our survey data you will see that too. If you look at between 05 that was really done in August of of and this March of 07 you'll see in just that eighteen month time frame the numbers have grown dramatically as a lot of people are at least looking at this.
5) PP: And is it No Child Left Behind or is it something else?
6) DP: No Child Left Behind is one, a huge issue, I mean there is no doubt about that. RTI, which is the Response To Intervention for Special Education is becoming a major issue in fact we are going to have a three day dedicated camp this summer for our schools, hopefully. The focus is on how the warehouse and how the RTI requirements and how we will track them and things like what we'll do so just that issue alone.
7) PP: Has caused you to start to looking at data more?
8) DP: Well it's it's forced us to make sure we design, because we have the luxury of designing these from scratch, of how do we want that set up and what kind of data do you need. And the other piece is what data makes sense. One of our biggest attractions to the current product is that we have the ability to put some rules in place that don't allow you to make stupid comparisons. Pardon the language. But, Johnny's shoe size doesn't have anything to do with his MEAP achievement.
9) PP: So you want it to be really quality controlled?
10) DP: And it is -- pretty rigidly
We have a uniform county-wide set of initial um demographic fields and some agreement on the parademographic fields as well across the whole all eight schools districts. And then they are allowed to customize beyond that to a certain degree for local y'know assessments that are unique to them but y'know. It's with the GLCE now for the curriculum thing there should be a lot more overlap now and less uniqueness.

[50] In fact, Don Pulte commented in a discussion not shown that Tichiochi ISD chose not to use MI-Tracker for just this reason.

The third area I consider (fifth in Mahajan, et al.'s list) is the bundling of options and product dependencies. Some products produced by textbook publishers and those with other infrastructural relationships within the state, for example through textbook contracts or other services, could have product relationships (ex: tests that are aligned with the book) that provide competitive advantage. Fourth (seventh in original) is market intervention. While the authors of one review study focused on commercial product interventions (e.g., patent violations), in this study NCLB and other federal requirements can be seen as factors that drive the diffusion process. Here also, potential advantage could accrue to larger firms with more resources to anticipate policy changes and to more rapidly modify their product than firms like Sucessline, Inc. And fifth (tenth in original list) is the rate of acceptance. This issue is at the heart of a discussion that has occurred in this dissertation and that will be discussed again at the end of this chapter: Who actually uses assessment information and for what purposes? Once the information is available, how deeply does it penetrate into organizational routines and individual practice? The relationship between features of systems and their actual use could be important to finding answers. Those systems like MI-Tracker, that are easier to use could be more quickly adopted, but only meet a small portion of requirements and become less competitive over time. All five of these issues illustrate the importance of a robust theoretical model to divining the directions this diffusion process may take.

The complex of issues related to data warehousing, including the relationships between ISDs and districts; the capabilities of a given software product; the relationship of policy instruments, including both NCLB and special education laws, and the role of the systems designers/implementers to set rules that structure what is permissible within the system are further illustrated by Don Pulte in Transcript 9.8. This chapter began with a discussion of MI-Tracker, one of the early and most popular products in the state. However, in Transcript 9.8, as in part of the chapter that preceded this one, it is possible to see that there could be successive generations of these software tools and that the process could involve restrictions on capabilities instituted by influential intermediaries.

This discussion ended by Don making important points about potential future relationships that could occur between ISDs and districts around these systems (turns 12-14). As ISDs consolidate to become the organizations that house and manage data for many districts, the ISDs will be in a position to define architectural constraints on the information systems. Beginning with the initial product selection and continuing through the modifications and expansions they approve, these ISD-based systems will have the potential to constrain information practices at the local level by either including or not including various data components. When Don indicates less uniqueness (turn 14), he highlights how these larger shared-data systems may enforce a uniformity across local contexts that could have political implications (Winner, 1986).

Practitioners

If we look to the literature that was reviewed in Chapter 2, much of it discussed the processes through which educators in different roles were integrating information into their practices. Phrases like *assessment literacy* (Boudett, City, and Murnane, 2005) and *cycles of data-informed inquiry* (Knapp, Copeland, & Swinnerton, 2007) remind us that the integration of information into an educational culture does not happen overnight. It includes steps and stages, and leadership is critical for the success of those efforts where individual classroom teachers may need to integrate new beliefs and approaches into their work.

This study shows that for some middle school science teachers, alignment of instruction to state standards is still an important issue. It has also included some glimpses into science teacher characteristics that could be factors in practice evolutions.

The seven portraits of practice in Chapter 5 illustrated the diversity of middle school science teacher practice in Michigan. And the organizational discussion in Chapter 7 showed diversity at the school and organizational level in Michigan, as well. In addressing the question of practitioner response to a continued diffusion of assessment practices and infrastructures into science education, I will use four different perspectives. First, I look at instructional alignment. This is one of the most fundamental aspects of an accountability system. Second, I look briefly at the broadening of the instructional agenda for educators. Not only do they need to teach specific concepts as mandated by public standards, they are required to teach all types of students. In these first two discussions I draw on both the Hardy Middle School case and a focus group interview based on a MEAP review meeting I observed. Third, I take a building and teamwork perspective and look at relationship dynamics within the schools. Here I use both the Hardy Middle School case and reports by Nancy Newman and Don Pulte. Finally, I discuss issues of classroom innovation, where science teachers, both as case study participants and as respondents to the MSTA survey, indicated they used classroom technologies that are compatible with a future era where ubiquitous technologies support systemic instruction (Bennett, 2002).

Analysis of Teacher Dispositions to Alignment

Undergirding the discussions that close this chapter are fundamental questions about science educators in the age of accountability: Which ones, and with what characteristics, will be likely to align their practice to systemic goals? Which ones resist, and why? After they concluded their work with reviewing potential MEAP items, the committee pictured in Photo 8.1, on the suggestion of Bobby Black of OEAA, acted as a focus group for me and answered questions about their backgrounds and the process of being reviewers for the state. This was a very senior collection of individuals with an average of 26 years' experience as science educators. I asked, as shown in Transcript 9.9, what they saw as common factors and patterns with teachers who resisted accountability and aligning instruction. Their response focused on resistance.

Transcript 9.9 - Focus group discusses teachers and accountability

1) PP: … have you seen any patterns in terms of teachers that resist alignment of instruction and using the results? Are there any patterns?
2) R2: I can't say that I've noticed a pattern in my past experience. It just seems that certain teachers .. um like to remain independent.. um feel that they can teach things in their own way better, not necessarily need to worry about these are really the objectives that need to be accomplished.. feel that there are other things that are more .. important to them whether for personal reasons or wherever that is coming from that .. feel need to be covered so they like to try to deviate from a curriculum.
3) R5: The mine, the MySpace people. In my experience they are the same people who don't want anyone to tell them anything about what they need to be doing or what is going on in their classroom. Y'know they want to be completely autonomous and .. uh usually I those are some of the older teachers and that's not to paint.

Transcript 9.9 - Focus group discusses teachers and accountability	
4)	R2: But there are some new ones too that are that way too.
5)	R5: Uhhuh that's true
6)	R2: That come out and say this is what I teach, this is what I am passionate about and I don't really care about this (pointing to document).
7)	PP: And they may be good teachers?
8)	R5: And because they say its important is
9)	R2: Oh yes, I'm not saying anything about whether they're not good teachers. It is that some will resist .. this type of thing (pointing to standards document) 'cause .. I can't give a reason for it.
10)	R4: Because they can't spend as much time on the things that they want to do with their classrooms. They don't resist assessment. They still will assess their students and they assess them well and some of them are **excellent teachers**. But, maybe they want to spend half a year on chemistry and not at all on light.
11)	PP: And its a personal choice?
12)	R4: Well, it shouldn't be.
13)	R3: Or they don't want to spend so much time on boring or worrying about that than.
14)	R4: Only cause.. many times to be honest, some of our older teachers have said "Been there, done that. It's going around again and again."
15)	R5: Yeah, that's true
16)	R4: That's pretty much the the consensus. They've done this enough enough to see it [new state standards] go around four times
17)	R5: Yeah I agree. They are asking us to change this and then we're only gonna change it and then they will have us do something else different.
18)	R4: And they get so frustrated cause we change it and then we go back to what we did before and they've done it so many times that they are just burned out.

The group indicated that while age could in some cases be seen as a common factor for those that resist, it was not determinative. While they said that some of the teachers who resisted aligning instruction were older, they indicated that there were some young ones also (turns 3-4).

Don Pulte also discussed equipment issues and emphasized the importance of alignment, as Transcript 9.10 shows.

Transcript 9.10 - Don Pulte discusses standards
Once they have a real specific [set of goals that say], hey at second grade you are supposed to cover this. And that will help us not cover for example dinosaurs at every single grade because everybody has to have a dinosaur unit because the kids like it. Well, no you don't actually because somebody else is doing it. And, here's your introductory and here's your responsibility and here's the mastery level. And our hope at least and we have talked about this with our schools is that it gets them to give some focus to science instruction and gets around the y'know I want to teach what I want to teach about science because I think it generally fits this broad category.

This perspective was reinforced by Carmen and Dianne at Hardy. In Transcript 9.11 (turn 5), they equate a negative and hostile mindset as factors in teacher resistance to what the focus group portrayed as more a matter of personal style or approach. They also discuss, in Transcript 9.11, how the MEAP results were used to identify instructional breakdowns to help in aligning teachers to state standards in the future at Hardy.

Transcript 9.11 - Dianne VanderMiller, Carmen DuRonder discuss alignment

1) DV: And you know the other piece that's been wonderful and especially for seventh grade science, which I know that's right where your area is, is that electricity is so prevalent. And of the four people sitting in that room three of them skipped that section of their instruction this year because they didn't have time. Now that means is that the eighth grade kids that took it, of course they didn't do well because those same people sitting at the table didn't teach it last year. So they were like
2) CD: At the same time we've never had a discussion about it. We've never broken down the test and looked at the test and said OK 40% of this is on electricity and 20% is on this 5% is on this, whatever. And so if you're going to skip something that's 40%. And so changing .. the mindset people think they need to teach what's in their book when necessarily they should be teaching what's on their GLCEs and what percentage that should be covered. Is it teaching to the test? No, it's your benchmarks.
3) DV: And I think there's also another piece of that is that the way that this has been handled. And when we went through and those four people spoke up uh those three peoples spoke up and said I didn't even teach electricity last year.
4) PP: uhuh
5) DV: I said OK what are gonna do? It wasn't Carmen these people didn't teach their stuff
6) CD: It's about how you show leadership.
7) PP: It's like what's the fix? It's to teach electricity.
8) DV: It's like what's the plan. Now they know and it's not a scold thing, it's an awareness. And and you.. We're not saying you didn't do your job or anything.
9) CD: It's an awareness. Well, it wasn't done so how are you gonna do it this year?
10) DV: But, you better bet that next year they are not skipping electricity because it's important. And its not something they should be skipping.

Hardy provides an example of school leaders taking an active role in the alignment process. The leader, in this case, was Dianne, who was an instructional coach and then was later promoted to assistant principal. This is a level of involvement that goes beyond what either Burt Wainwright in Swallow or Stan Dubovski in Avon Falls did, because Dianne was engaged not only in general requirements for teachers to align to the standards, but also in the specific topical areas and logistics the teachers faced. Not shown in any of these transcripts are some comments that Dianne and Carmen made that indicated part of the problem those science teachers may have faced was that the kit schedule was not adhered to, and part of the fix was to make sure the kits arrived to the school on time.

Broadening of the Instructional Agenda

Not only are science educators in Michigan being pulled towards a specific set of instructional goals (the GLCEs), they are also being pushed to teach for a wider range of students than in the past. Throughout this chapter, special education and issues of disability have taken a greater role than in earlier parts of the dissertation. and Dianne VanderMiller's comments in Transcript 9.3 highlight the issues that practitioners at Hardy were facing when they equated disadvantaged subgroups with those requiring special instructional accommodations, as Diane had. Not only are there often strong correlations between these subgroups (Wade and Zone, 2000), but in Dianne's view the instructional challenge was the same: finding ways to teach students who do not fit the typical profile of accomplished students. Transcript 9.12 is a continuation of the earlier discussion where she makes clear her position (Transcript 9.3, turn 2) that teachers at Hardy have a greater responsibility and that NCLB provides the leverage to make that point clear to them.

Transcript 9.12 - Diane VanderMiller discusses NCLB and its role in focusing Special Ed services (continued from transcript 9.2).
16) PP: So this is really a side kind of a side effect of NCLB that? is it a good thing?
17) DV: I think finally we have teeth that lets us say you have to service those kids better. You have to teach at their grade level. Your job as a Special Ed educator, your job is to teach them subjects that may be difficult but use the accommodations that you've learned through your degree to access that knowledge for them. **Find a way.** And I'm not sure that's a whole lot different than finding a way .. to engage a disenfranchised kid y'know or a kid whose got a million other horrific things going on in their life outside of school that we can't even imagine in our lives.
18) PP: Right. Right.
19) DV: And they come to school and work for certain teachers. And why is that? Those teachers have strategies with those types of kids. We've got to find a way to do that with our other subgroups.

This same issue arises again in the discussion with the MEAP focus group. In Transcript 9.13, they discuss NCLB and segue into the issue of special education students. One member (R5) believes some of his students should be sent to a special class in an ISD (turn 8). Unlike the special science schools that Pete Darmond and Corrine Eaton discussed, this ISD school would not be for advanced students. Rather, it would be the type of special education tracking that recent federal legislation, including NCLB and the Individuals with Disabilities Act (IDEA), highlight in requiring that students be placed in mainstream setting (i.e., least restrictive environment) when possible. Don Pulte provided a congruent perspective earlier in Transcript 9.8 (turn 10) where he references Response-to-Intervention, which is part of IDEA, as a motivator along with NCLB for the push to use data for instructional purposes.

Transcript 9.13 - Focus group discusses NCLB (continued from Transcript 9.9)

19) PP: But has No Child Left Behind made it different?
20) R5: Made it more difficult
21) R4: Made it more accountable
22) R5: Because we're um well
23) R4: I have an accountability to my principal because I'm a department head so my teachers and I have a teacher that does that. And .. um he's doing better but its been very difficult. You have to be quite a diplomat. You have to make sure cause he's a phenomenal teacher. But he's seen it happen so many times.
24) R5: I have an issue with adequate yearly progress because and its gonna by 2014 is it or 1s is it? 100% of the kids are passing. Uhm come on? It's not gonna happen.
25) R4: You can always figure that one child in the classroom or five or ten that
26) R5: I've got some mainstreamed kids that are supposed to be at an ISD or should be and are not and y'know we did 91% and we're proud of that and that included our special ed population which is which is a pretty good population and we'll make adequate yearly progress based on that but then in future years, we're gonna fail. We'll be failing in science because we're only getting 92% and that I have a problem with.
27) R4: What an awesome score that is.
28) R2: Yea.
29) R5: We're happy

Across this dissertation we see issues of policy, student ability and cognition, technology, and administration and teamwork coming together again and again in participant discussions that may initially begin with instructional alignment (what gets taught) and then broaden to include different groups of students (who gets taught). The two themes of standardization of instruction and universal education are two themes emblematic of the complex period this study sampled.

Evolving Team Relations

The progression of this dissertation – from classrooms to buildings – highlighted by the signs used to begin Chapter 5 and this chapter is emblematic of the realignment of instructional responsibility from individual teachers to broader groups that was highlighted in much of the literature reviewed in Chapter 2.

This study has produced some glimpses of the transitions that practitioners involved in science education had been going through during this time. In Transcripts 9.9 and 9.11, the MEAP committee/focus group and Don Pulte described how some teachers were inclined to teach what they wanted rather than aligning to common standards. Chapter 8 showed nascent discussions about how to unify science instruction -- across grades at Swallow and across adjacent classrooms, both teaching science to seventh graders, at Avon Falls. At Hardy there were similar discussions, including the group of four teachers discussing how electricity was or was not being taught (Transcript 9.11). Beyond this issue of science teachers' individuality, other processes were occurring within the scope of this study where roles were being realigned in ways that are parallel to the availability of instructionally-relevant information and in ways consistent with this broadening of instructional responsibility.

Transcript 9.14 - Don Pulte discusses teachers beginning to face accountability.

1) PP: With your Tichiochi hat on um and being involved with teachers and principals, what's the cultural issue or issues that you're seeing if this information were to be used for instructional purposes. Who's buying in, who's not. Is there any pattern you see?

2) DP:No from my take, and if you've worked in technology you can appreciate this, I don't think its any different from that old third, third, third. Y'know a third of them are kinda gung ho for anything technology related, a third are like let's wait and see, and a third are like hell no we're not going to do this.

3) PP: OK

4) DP: Um I don't know if have any magic bullet with them. What I think the potential barriers that we run into is that when you begin to compare classrooms to each other that teachers have never had that level of scrutiny probably at a, or accountability I'll be honest with you that's what it is, accountability at that level

Transcript 9.14 - Don Pulte discusses teachers beginning to face accountability.
before, where you could literally take and hang up, anonymously, five classrooms on the wall next to each other and all of a sudden you wonder why classroom C is doing so much .. worse on multiple achievement measures that the others or whatever if the scores aren't what they should be. I think that's a .. that's **a big issue for teachers**.

Don Pulte, from his position as an associate superintendent for the Tichiochi ISD, was able to comment on some of the challenges that teachers had been facing during this period. His description in Transcripts 9.14 is similar to his discussion in Transcript 9.8, in that it integrates a variety of perspectives including policy, technology, classrooms, and instructional processes. This mixture parallels the broad resources used in the emerging literature on these processes shown earlier in Table 2.9. Don illustrates the transitional nature of this period; as he talks about beginning to compare different classrooms (turn 4), he is generally discussing an evolutionary process enabled by information. It is reasonable to infer that with continued pressure from policy in NCLB, IDEA, or successors to that legislation, these practices would continue to become more institutionalized.

Nancy Newman, from her position as an ISD curriculum administrator, also saw many teachers beginning to look at data in forums where questions are asked about their practice. Most of Nancy's experience has been with data being used to inform literacy teachers, and what she had seen is that some teachers had more difficulty than others in participating in forums where their practice was discussed using student performance data. In Transcript 9.15, she explains that the issues with using data are essentially issues about being a reflective teacher and supporting a climate of practice transparency (Little, 2002).

As described in this chapter's keystone case study, the science team at Hardy did meet with Dianne VanderMiller and discussed their MEAP results, as discussed earlier in Transcript 9.3. The situation at Hardy, however, is more contentious. As with all of the other reports where participants were asked about which teachers are likely to be partners in accountable education and which ones are likely to resist, Carmen DuRonder clearly stated that it was not a matter of the young or the old resisting, but a matter of the teachers' commitment to students. In Transcript 9.3, Carmen and Dianne described the transition that was occurring within their school. Carmen sees herself as the reformer (turn 1) that will show the teachers the way forward with accountable practice. She and Dianne highlight that there is not only a process of growing acceptance taking place, but also that many teachers supportive of the reforms were coming out of the woodwork and being more vocal about the need to be committed to educating students, rather than just doing a job (turns 1-3).

Transcript 9.15 - Nancy Newman discusses challenges teachers face in being reflective, using data.

1) PP: Do you see some teachers who don't get it?
2) NN: Oh well,
3) PP: Is it a big problem?
4) NN: For me, the issue is whether you can reflect on your practice.
5) PP: OK.
6) NN: OK and to do that effectively, you have to be able to look... just take a step back and say what am I seeing here? What is the evidence?
7) PP: OK.
8) NN: Am I am I being quick to judge?
9) PP: Judge the kids? Judge yourself?
10) NN: Yeah. Whatever. Just judge. And I would say that's overwhelmingly what happens to teachers they are quick to judge and they have a hard time stepping back and doing the reflection they need to meet a student's need. So when I say someone is having a hard time with data, for me it that means they are having a hard time reflecting with the data.
11) PP: OK.
12) NN: And when you have a really.. either because A they don't understand what it means, which we can usually get past. I don't know of anybody who can't eventually with support understand what a dataset is or what it is saying. I think we can get past that. But it's sort of that piece that's about going to the next link. That's and when I see teachers that can't get that I get scared I just get really scared because I wonder how are they ever going to improve their practice if they can't do that work?

These different accounts from school and district leaders, and many more not included in the interests of space, portray the processes in schools as not being simply addressed by time, as older teachers retire and new teachers enter. In fact, they all reinforce the position that, without conscious and deliberate leadership work, many schools could have had practitioners who would not naturally develop the reflective habits important for being able to use the abundance of data that the developing infrastructures seem poised to provide.

In the discussion in Transcript 9.16 is another integration of topics that speaks to the transitional nature of the period for Hardy Middle School. Ann and Dianne describe the tensions within the staff between those who have created a climate of opposition to reform at the expense of the children's education and those who, like Ann and Dianne, wanted to prioritize education and supported accountability as a way to strengthen, rather than diminish, the team of educators in the school.

Transcript 9.16 - Dianne VanderMiller and Ann Pobzerniac discuss the transitional time in Hardy Middle School in terms of reflection, alignment, teamwork, and teacher voice.

1) PP Where do you see those who are resistant and where do you see those embracing um more data driven more reflective approach. Is it more about reflection? Reflecting on their practice?

2) AP: The teachers that I know who have struggled with it, I think it is about reflection. What I think..they want to do what they want to do.

3) DV: Nobody's going to tell me how to breathe. I mean that's a quote from this year. No one's gonna tell me how to breathe.

4) PP: How to breathe?

5) DV: Yeah. and I said the state already did. By saying here's grade level content expectations. And I can help you [inaudible] do that. I've never done that before and I will say there was a marginal group of teachers in the middle .. that are now so empowered. And there is a very positive undercurrent of voice of people that are .. not being bullied anymore.

6) PP: And when you say the middle, what is the middle the middle of what?

7) DV: Attitude I would say.

8) PP: Sitting on the fence?

9) AP: Well, let me put it this way. there was a group of teachers and unfortunately I was .. the one who like always like y'know what I wanna make a change. So I was always the kindof voice of that little group. And we all kindof hang out. So its like we would always like meet every other Friday at the bar, and

10) DV: What can we do..

11) AP: And kindof complain about these things why can't we do this and why can't we do this why can't we do this. But none of us would want to take a stand.

12) DV: Because like did she tell you the dynamic

13) AP: I didn't start it till last year. And then I took a stand and said y'know what I gonna join the school improvement team at least I can get at least I have a voice in there. And I know what there's nine of us. I can bring our nine voices to this group and hopefully make some kindof change. Because we all we all want to do the best things for kids. But we were there to do a job. For us, we're there to do a job we're going to do our job. But there were those who were so loud .. constantly.. that kindof pushed us down.

14) DV: And it doesn't even have to be about curriculum. It can be about policy or parking for god's sake or ice in the parking lot it's the same issues all the time. But I'll tell you what's been very powerful is that the pocket of nine realizes that there's this pocket of nine over here and this pocket of nine over here and pretty and now you have five pockets of nine. And, we all want to do the right things for kids and now

Transcript 9.16 - Dianne VanderMiller and Ann Pobzerniac discuss the transitional time in Hardy Middle School in terms of reflection, alignment, teamwork, and teacher voice.

they have a voice.

Evidence of Innovation by Science Educators

The classroom teacher is in a gate keeping position regarding her classroom practice. This study is not alone in finding that teachers are resistant to innovation. In fact, many of the bifurcated models of education look at the technical core of the educational process as a location of resistance, the place where good ideas riding on the sled of policy hit the bare patch of exposed asphalt and fail as fast as they came. Changing classroom practice is hard, so paths that may penetrate classroom practice are important to understand. Any innovations that may improve the odds of performance are important to understand.

One way to consider innovation with science teachers in this study is to look at their use of classroom assessment technology. Faith Churchill and Jonathan Brunson had both begun to use classroom scanners in the year that I studied them and reported little difficulty in continuing to do so. Christy Connolly had begun using Moodle. Across these three teachers are varying levels of experience, from Jonathan, just a few years into his career, to Christy in her twelfth and Faith in her thirtieth. In the MSTA survey, teachers were asked if they used any special technology in their classrooms that supported their assessing students. Forty percent of all teacher respondents, and one-third of middle school teachers, reported using something to support their assessment practices, as illustrated by Figure 9.2. These responses suggest that many science teachers in Michigan are beginning to use various classroom assessment technologies and that the teachers profiled in this dissertation are not unique. Furthermore, since elementary school teachers often teach multiple subjects, it is conceivable that in responding to this question those teachers also included assessment technologies that they use for other subjects, as the commercial support for other subjects, specifically literacy, has been documented in Chapter 2.

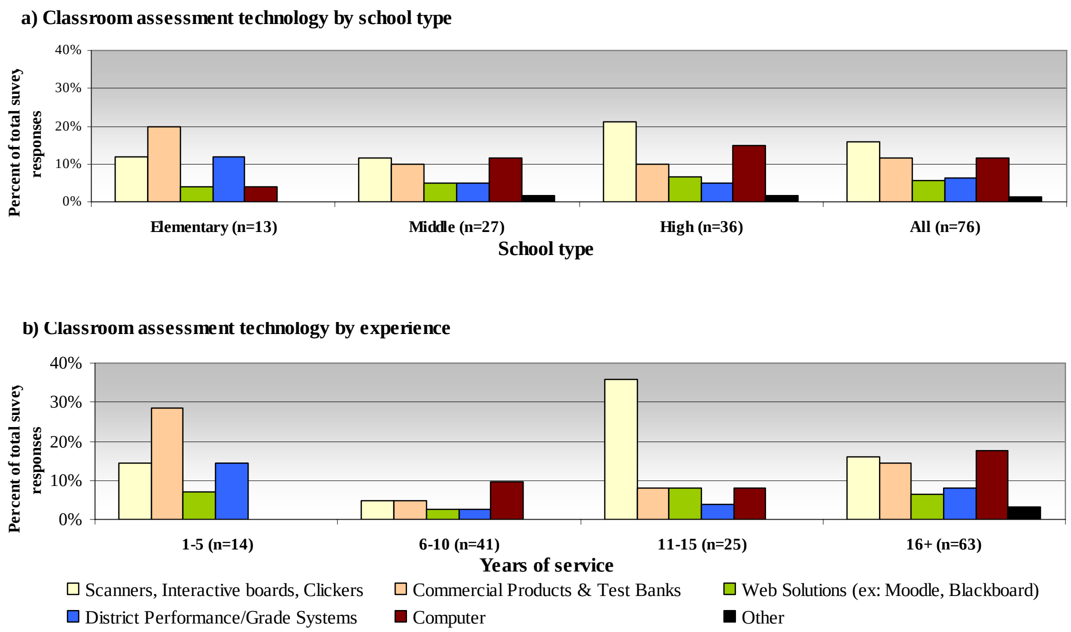


Figure 9.2 - MSTA survey responses for classroom assessment technology.

This small set of responses further reinforces the conclusion that age is not a determinate factor and that many experienced teachers are engaging in new classroom techniques that can support future informational infrastructures for student performance. Many science teachers, including some like Faith Churchill, who have significant experience in the classroom seem as engaged with multiple technologies as newer teachers. In considering this period of evolution in both information infrastructures and systemic practices, the innovativeness of science teachers, especially experienced ones, may be a valuable resource in helping science education becoming a more publicly accountable practice.

Especially About the Future

This chapter both concludes the empirical core of the dissertation and introduces questions about what may be next for science educators in case study contexts and, indeed, in the State of Michigan in the area of science assessment. In taking on the issue of historical change, I opened up a new set of inquiries and drew upon all parts of this study's evidentiary base. What has emerged is a sketch that highlights the interplay between practice and technology, between organization and infrastructure. If we consider the two data warehousing surveys included in this discussion as the beginning of a process whereby technologies for representing students and their performance will eventually cover the state, it is also possible to speculate that this process can include shifts in practice from an era of individualized teaching to shared responsibility.

This chapter, as the previous two have, shows an integration of topics across cases of very different types. By crossing from individuals in schools to those in ISDs and then drawing on surveys, a portrait of the historical period emerges with topical coherence from different perspectives. Seeing the previous few years as a time of rapid take-off, seeing this period of accountability as being largely fueled by federal legislation, and seeing the tensions that these changes are raising in practice are all reasons to envision broad movements similar to what Kuhn (1970) had referred to in the scientific communities as paradigms. Similarly, the relationship of this period to new technologies for collecting and storing information is similar to the paradigmatic description provided by Bowker (2006), also in a study of natural scientists, that points to changes in what counts as evidence. In Bowker's description, and I would argue across this dissertation, the general ability of the database to store and distribute information relevant to many contexts is instrumental in supporting realignment of those practices. As Bohr has reminded us, however, the inferential process whether about what has occurred or what will occur is difficult business.

It is possible to even construct these shifts in paradigmatic terms, where the portraits of practice presented earlier in Chapter 5 become representatives of an era before all science teachers used a coherent set of benchmarks to structure their instruction. It is possible to conceive of this new paradigm as one where school leaders are able, with the help of common curricular constructs like learning progressions, to participate as instructional leaders across grades for science (Piety, 2007.) It is possible to further speculate about how the designs of data management systems like multi-ISD data warehouses and DATA4SS can, through their information architecture (Piety and Palincsar, 2006), allow representations of student cognition to be codified into computational supports.

Much about the next, even immediate, steps for Michigan educators concerned with science education remains unclear. It is not clear from the evidence in this study whether there were aspects of this historical tension between teaching as a closed door activity and teaching as a transparent and communally accountable practice that related specifically to science. Because of the segmented nature of science and the multitude of instruction options available to science teachers, there is a possibility that topical alignment was one of the most important dimensions for many practitioners. Much of the evidence available to this dissertation suggests that topical alignment remained an important challenge within Michigan during the time of this study.

Chapter 10 Conclusion

"Synergy means behavior of whole systems unpredicted by the behavior of their parts taken separately." ~ R. Buckminster Fuller

This study has produced an empirically-developed depiction of educational practice – a systemic model – using assessments of student learning as a primary organizing construct. This model can be compared to the charts that greet subway travelers in many cities like Washington, D.C. These charts use a simplified map with different colored lines for trains, symbols for stations, and a table showing fares at different times of the day. Taken together, this assemblage can be thought of as an abstracted systemic model that highlights information most important for riders of the train. It helps in trip planning and shows limitations in the system where the train may not be the best method of transport for some trips, and it does so efficiently by using a simplified and abstracted set of images. The systemic model in this study similarly abstracts a more complex reality by highlighting certain elements. These elements are largely derived from the structure of the educational policy that makes student assessment information both a school outcome variable and also a resource for educational decisions.

In the natural sciences, models abound. Ecological models may illustrate the possible effects of pesticides; models from fossil records help support the theory of evolution. As a product located within a specific scientific discourse(s), a model may be useful in advancing a field or focusing internal discussions onto important questions. A model, then, can be seen as localized to its empirical foundation and also as a text with mediational potential within and across discourse communities. As a scientific model, this study's products can be evaluated in several ways. One way is in terms of evidentiary quantity and quality. The model can be evaluated based on what is included or excluded, the depth at which different elements were covered, and other intrinsic characteristics of its evidentiary components. It can also be evaluated in terms of the description it provides. Is that description useful, and for what purposes? And it can be looked at in terms of dialogue conversations, as to which discourses the model may support. The primary contribution of this study, then, is a model of educational practice and how that model can be used for science education and, more broadly, for educational reform.

Because the model is one of many possible models that could be produced using differing methods or in different contexts, the results of this research will be first summarized in terms of key features of the model. While these features are related to the research questions that helped introduce the study, they may also be applied to other contexts and questions. Because these features show different facets of systemic character, this first section is titled "Evidence of System." After discussing these main features, I then reflect on the nature of the evidence in the model in terms of its limitations, coherence, and congruence with other evidence. This evidentiary discussion helps to transition to two discussions on using the model to reason about policy and practice. The policy discussion uses the model in two ways: to describe conditions where policy is to be implemented and to ask whether the model of education implicit in a policy fits this model empirically derived from indigenous practice. The practice discussion uses the model to describe an expanded form of professional development that includes not only teachers and their classroom work, but also boundary practices and professionals who can benefit from exploiting these resources for synergy. I then discuss how this model can contribute to the nascent literature on systemic uses of information in education I reviewed in Chapter 2, as well as to theories of educational organizations and literatures that conceive of how organizations connect through individuals (brokers/boundary spanners), artifacts (boundary objects/mobiles), and practices I reviewed in Chapter 3. I conclude with a historical discussion that places this research into a larger narrative of historical change and identifies future research steps within a description of educational systems operating under greater expectations for, and reliance upon, information infrastructures and practices.

Evidence of System: Five Key Model Features

In the introduction to this dissertation, I described three different ways that the term "system" could be used in this study: as a set of things or artifacts (and the individuals who use the artifacts); as a set of social processes (which may include artifacts), and as reciprocal relations where synergy is evident. The term *evidence of system* is a phrase that can be applied to all of these ways. This study shows evidence of system through three main model features:

1. **Two dominant types of assessment systems: individual teacher classroom practice and the state's annual accountability test.**

2. **Support for *both hierarchical and network* depictions of educational organizational practice.**

3. ***Boundary practices* that exhibit a range of important systemic issues and opportunities for alignment and synergy.**

Further, the model has two other general features: variation and dynamic qualities. Rather than showing these systemic characteristics in static and fixed terms, the model shows them as diverse, evolving and contingent:

4. **The model illustrates practice variation and variable interconnection.**

5. **The system was in a period of change due to rapid diffusion of information infrastructures and concomitant pressures to reshape practice toward collaborative information-driven activities.**

I will discuss aspects of each of these features below before transitioning to a discussion of some aspects of the evidence base and how this model can be used for policy and practice.

Dual Systems and Classroom System Independence

The model contains two principal types or species of assessment systems operating within the state of Michigan during the time of the study. The first was embedded in classroom practice to support instruction, grading, and occasionally communication with parents. The second type was the annual accountability test. While classroom systems were largely closed, the state system crossed contexts and linked the state to individual schools and teachers.

This model suggests that in many cases, seventh-grade science classroom assessment practice and the state test remained largely independent activities. Across all seven teachers in this study, including those in the same school as other case-study teachers, classroom assessment practices were essentially teacher-specific. Not only were most of the seventh-grade classroom teachers disconnected from state standards and showed little effect of the annual test in their work, but their practices were not aligned with their fellow teachers' practices. These classroom practices for science education were also largely disconnected from the teachers' leadership. Furthermore, little evidence was found that alternative assessment systems were in use. While the possibility could not be eliminated that other, intermediate assessment systems for seventh-grade science operated across schools or districts, it is likely that if alternative systems had existed anywhere in Michigan during the time of this study, their use was localized and not well publicized. This feature contrasts seventh-grade science education with other subject areas, including early literacy, that had multiple assessment systems available to practitioners, as discussed in the studies reviewed in Chapter 2.

Both Hierarchical and Network Characteristics

This model has both hierarchical and network dimensions. At the local levels, the common nesting of classrooms within schools within districts was hierarchical. As the structures moved further from classrooms into the Intermediate School District (ISD) and Math/Science Center layers, the character of the organizations became more orthogonal, with the relationship of schools to intermediate organizations more likely to be many-to-many and the services a given teacher might use coming from more than one intermediate location.

The model contained a web of participation that included other types of practitioners in schools, districts, and support centers. The web was used to triangulate on state-level processes and issues related to science education. In terms of professional communities, the study included the professional associations for science education, assessment, and administration as additional cases. These are dues-based communities that individuals may join to pursue common interests, rather than general automatic categories. As cases, they were then related to individuals and other organizations in the study. While the administration and testing associations had overlaps in terms of individuals and events, the science educators' association was largely separate from those two. This professional community independence paralleled the independence of science educators at the school level. These state-level organizations provided another way for individuals to be grouped in addition to their formal employers, indicating a network that can contain more than one type of hierarchy.

Systemic Opportunities in State-system Boundary Practices

This model shows two important mechanisms for systemic connection. Both were related to the state's annual accountability test. The first is science educator participation in the test development process managed by the state government. Participating in the development and review of the state test items was reported to strengthen understanding of assessment practices and commitment to aligning instruction to common standards. The second mechanism occurred within local schools as educators and leaders reviewed their annual test results and identified specific instructional gaps. In two cases (Swallow and Hardy), the state results identified learning targets in the state standards that were not being taught by case study teachers. In another case (Avon Falls), the results showed that specific terminology used in state standards was important to understand in order to index the depth at which the material would be tested. In Avon Falls the results evidenced that the teachers understood the concepts listed in the standard, but not the way that test developers interpreted those terms or the depth of understanding that the test would assess.

Analysis of conversations about assessment items, both in the state government's development process and in schools as they reviewed their results, indicated that broad systemic topics emerged around discussions of the items. These topics, which arose in both state and local contexts, included state-level matters of standards, school-level matters of teaching methods and classroom equipment, and student issues of cognition and belief systems (misconceptions). From a systemic view, despite the boundary practices that this study has identified, the state assessment system and the local classroom systems in this model appear out of alignment. While some alignments are topical (as in the case of content standards), others are more difficult, as they relate to issues of classroom practices and laboratory equipment, including science kits that complicate alignment. Importantly, while many schools appeared to review the state test results, the connection between these reviews and instructional practice was anecdotal and uncertain. I was able to uncover no examples of classroom practice in either instruction or assessment being deliberately adapted to meet the needs of the standards, although some teachers expressed support for the idea.

Variation in Practice

Across the domain studied, variation in practice and interconnection was observed. For example, the depiction of the system as involving a largely independent classroom layer is not complete and universal. Some teachers were more connected to systemic components, and their practices were more consistent with principles of accountability than others. The relationships between school leaders varied, as well. While none of the five cases featured school principals acting as instructional leaders for science education, across the five there were examples of interest and engagement in science instruction. The principal at Crimson Middle School, Ed Bedminster, helped kids finish science projects, and two school leaders (Wainwright in Swallow, and VanderMiller in Hardy) reviewed results of the annual test with their science teachers. There was also variation in the reform epistemology of participants across this study. While some teachers took a position contrary to contemporary accountability policies such as NCLB, others held views that were strongly supportive. The only practice dimension that was related to these teachers' positions was the degree to which they were involved in their professional community. None of this variation, however, demonstrated classroom alignment to standards, leading to a largely bifurcated view of educational assessment practice.

The relationships between individual teachers and the broader systemic context varied significantly. Some teachers did not participate in state or national science teacher forums, and some did. Some did not participate in state-level activities involved with developing the annual test and some did, reporting significant benefits from the process.

Across the schools, leadership structure and style also varied. In three of the schools the study had enough opportunity for observation to support a sketch of their interpretation of the state test results, and each leader framed the issues surrounding the test in a different way. In one case (Swallow), the annual test was explicitly addressed and the response was task-oriented. In another case (Avon Falls), the school principal did not directly address the details of the state test, but rather developed a community response. In that case, the state test was perceived as a threat to the emotional comfort of the staff and the principal responded by launching a team-building exercise within the school. In the third case (Hardy), the school was both undergoing an administrative realignment to respond to low student performance and beginning to use a data-reporting system to help in analyzing and visualizing the annual results. That school's leadership team embraced the use of results as a management tool and also adopted a pro-accountability stance: all students could and would succeed on the state test, and student failure would be traced to professional failure of school leaders and staff.

Despite the variation across the cases, the annual state test results were seen as mediators of school interactions. These glimpses of practice suggest that the availability of public information about student performance in science can be a resource for strengthening practitioner relationships.

A Dynamic and Evolving System

These model properties need to be framed within the understanding that there was evidence that the educational organizations and individuals studied were in a period of evolution and change, and anticipating how these practitioners will be working several years later may require additional work. One of the key evidences of this evolutionary change was the widespread adoption of information infrastructures across the state. These infrastructures were often housed in districts and intermediate districts, and some were more comprehensive than others. Some were envisioned as data warehouses with the capability to integrate a variety of information, including multiple assessment systems. The state government was not only a provider of information to these projects, but also a partner in some. In one case it participated in a federal grant to create information standards for sharing of data that the State of Michigan received from local schools. This student information would be made available to all local schools so that students' records could follow them if they moved across districts.

This progression of information infrastructures across the state during the time of this study has important implications for the study's results. First, these infrastructures would be capable of supporting science assessment systems in the future, so the finding that the state system and the local classroom systems are largely distinct should be tempered by the potential for the inclusion of other assessment systems somewhere down the road. Perhaps other assessment systems could come from commercial providers or from within the schools or districts. A second aspect of this model is that many educational organizations were also in periods of transition. The ways they used student performance information were evolving, with new organizational routines being developed.

Reflecting on Evidentiary Limitations, Congruence, and Coherence

This model has some evidentiary limitations. While I can only begin to explore them, it is important to identify some of the significant constraints on this evidence before discussing how the model can be applied to various contexts.

One source of limitations exists in the types of participants. The case study schools and teachers represent rural locations with fewer minority students than are found in some parts of the state. Also, as discussed in Chapter 6, the experience level of the teachers in this study would seem to be higher than the state average. It is conceivable that there are many science teachers with less experience in Michigan and that their practices may be different than those represented in the model. Those differences are difficult to assess, however. This limited representation is also evident in the survey of members of Michigan's Science Teachers' Association (MSTA) discussed in Chapter 6. Since the survey did not have random or complete sampling of the population of teachers, but rather sampled mostly those members of the MSTA active enough to participate, it is used as a supplemental data source. Pursuing the nature of these limitations with greater precision is severely challenged by a lack of reliable information on teachers' experience in the state.[51]

There are reasons to consider Michigan's educational environment as similar in some ways to other states. In other ways, however, Michigan is unique. The broad alignment between the trends exhibited in the literature reviewed in Chapter 2 and the processes of technological diffusion occurring across Michigan reviewed in Chapter 9 suggest that what was occurring in terms of adoption of assessment-related infrastructures and practices of data use in Michigan was similar to transitions occurring in other states during this time. Some of the material connections that link schools with commercial concerns, including textbook publishers, software vendors, and test material producers reviewed in Chapters 7, are not specific to Michigan. Other states may be experiencing similar processes and have similar commercial entities operating with them. In terms of its test development process and testing cycle, however, Michigan's approach is not the norm. Not all states use practitioners in their test development process, and many states do not test and release results within the same school year, as Michigan does. Some test at the end of the year and release results in the summer. Also, Michigan has a large number of intermediate school districts, making it one of the most extreme states in terms of its organizational diversity.

Important limitations may also come from the fact the model was constructed from assessment information rather than from other potential mechanisms for systematicity. Also, concerns could be raised about the candor or transparency of participant claims, as prior research indicates that the use of information in school decisions is one area where participants tend to over-report (Coburn, 2004; Ikemoto and Marsh, 2007). The analysis of practice texts, documents used as part of regular work, was used to triangulate on specific claims, and many of the claims have been kept at a general level where the documentary evidence was strong to minimize issues of participant candor.

[51] Because a teacher can move from district to district, the systems such as those maintained by the state licensure office will not be reliable. This was confirmed in email conversation with Mr. Anthony Beal of the Michigan Teacher Licensing Office and a representative of the Michigan Center for Educational Performance & Information in December, 2007.

Despite some limitations in the model's participants and its scope, it is largely congruent with the general portrait of other educational contexts in the literature reviewed in Chapter 2 and is also internally coherent. The literature reviewed in Chapter 2 provided evidence of a recent surge in interest in information use across a variety of educational contexts. This literature placed greater emphasis on literacy and mathematics than other subjects, and often connected this interest to NCLB. The Chapter 2 literature showed studies that were often early pilot projects, rather than sources of information used broadly in daily practice. Some of the studies reviewed discussed barriers to data use, including the concerns of teachers, principals, and administrators regarding what counts as evidence and how systemic information is often not used to drive decision-making, but rather only in supportive or symbolic ways. This model's evidence is largely congruent with these depictions.

This model also exhibits several characteristics of internal coherence. In addition to elements being linked geographically in the state of Michigan and within the same time period, there are many cross-evidentiary connections. Chapter 6 and Chapter 7 highlighted relationships between participating individuals and institutions. Participants in this study were also related by their understanding of certain past events, such as the closure of a Math/Science Center that one participant described in Chapter 6 and another explained in Chapter 7. Different model entities were also linked by their association with specific materials and technology providers (Coburn, 2004), as shown in Chapter 7, and by their participation in state-level professional associations, as shown in Chapter 6, Chapter 7, and Chapter 9. Also, some of the participants in this study took part in the same actual events through joint participation in the state government's test development process. Other participants had personal/professional relationships with each other, and in some cases, relationships to the researcher and the University of Michigan.

Informing Policy and Practice

The model presented by this study can be used both to heighten understanding of how policy and practice are related and also to envision ways that practice can be strengthened and evolve to meet policy.

Dialectic with Policy: Multiple Levels of Adaptations

How does this model of science education practice inform policy? As with other scientific models in evolving social practices that serve as mediational elements to look at both what is and what might be, this model can be used as a way to evaluate the relationship of policy to practice. One type of evaluation uses the model as a description of the landscape over which a policy could be implemented. This evaluation includes issues of readiness for anticipated policy, such as science education becoming part of AYP. Another evaluation could use the model to evaluate the fit between the policy and the practice world the model describes. This evaluation uses the empirical model as a means to identify weaknesses and missing connections.

This study has shown that the next generation of science education policy might consider not only specific learning goals and how they can be worded, but also issues of classroom equipment and the variety of circumstances within which teachers operate.

McLaughlin (2008) frames the policy discussion in terms of macro and micro actors, saying, “Micro is not a little, bitty macro. Macro tools are technical; micro tools are adaptive” (p. 15). McLaughlin continues by describing the adaptive processes that local actors in communities of teachers go through. They are what Weatherley and Lipiski (1977) called the “Street Level Bureaucrats.” They are the final arbiters of policy. This model shows elements that are compatible with this view, as the state content standards can be seen as technical, macro instruments and the seven case-study teachers as micro-level actors whose work is adaptive. In this study, both the macro and micro are viewed with the same lens (Latour, 2005). One teacher in this study, whose failure to teach a science benchmark was highlighted in Chapter 8, for example, was shown to be out of alignment with macro expectations. At the same time, it was shown that his approach to teaching that material was related to the fact that he taught in a very small district and was responsible for both seventh and eighth-grade science, and thus his instructional choice was adaptively situated within his local circumstances.

This model allows this micro-macro policy discussion to be extended further. By including the intermediate structures – those formal educational organizations that include districts and support centers – and by looking at the state-level professional associations for science teaching, assessment/evaluation, and school administration, this study traces particular paths from macro to micro. In walking across these trails from individual classrooms to state-level functions and documenting particular individual relationships, different *types of adaptations* can be seen. McLaughlin continues:

> Macro and micro level actors have different "keyboards" at their disposal and at the level of practice, important elements include those outside the formal policy system—parent support, community engagement in education, local political economy, community based organizations and so on. (2008, p.16.)

This study included some individuals and organizations that existed between classrooms and the state and exhibited different types of adaptations that occur in this middle level. For example, in Chapter 6 a senior science specialist at one ISD discussed learning, at a meeting of state-wide science leaders, that a Math/Science Center in another part of the state had been eliminated. In Chapter 7, the ISD curriculum director who decided to take that function back from the science specialists explained her decision to prioritize central office cohesion over science expertise. In Chapter 9, an assistant superintendent from another ISD discussed implementing assessment systems and procedures that would require individual teachers to be compared to each other within the local schools. At another ISD, a leader in the assessment/evaluation community, *an opinion leader* (Rogers, 1995), was also an advocate for a particular commercial publisher's software product. Chapter 9 showed how that vendor's product increased in market share during the few years leading up to the time of this study. All of these intermediate-level actors have keyboards at their disposal, as did the school principals who showed examples of different local responses to the same annual test. The model shows how the levers these actors push and pull may be important to consider in theorizing about the types of transitions that can occur from macro and technical policy to micro and adaptive practice.

In looking from this research model back towards policy, this study shows how the definition of the educational system, the model, within NCLB is lean compared to the model this study produced. While NCLB describes students and teachers in schools and systems in fairly hierarchical relationships, this model shows other potentially important systemic components, such as professional associations and materials providers, that may provide varied and unequal resources to educators in an ecological view of education (Weaver-Hightower, 2008). Textbooks and instructional materials have the ability to influence practitioners' enactment and professional learning (Ball and Cohen, 1996; Davis and Krajcik, 2005). In addition to selling a range of products to local schools, this model showed these organizations' involvement at the administrative level in the emergence of data warehouse solutions and the leverage and influence they may be able to exert over science education as they provide both systems of measurement and resources for instruction.

This model shows how the conceptions of education in policy and practice can be reconciled. It demonstrates that, even if policy is constructed not to consider all aspects of practice in its instrumental form, a model such as this one can assist in the discourse about implementation considerations.

A Tool for Practice: Professional Development for Systemic Actors

What can practitioners learn from this study? One way to apply lessons from research to practice is through the use of interventions, often targeting teachers, called professional development (PD) events. The model this study produced suggests the lessons from this research can be applied more broadly than with teachers alone. For while this study found synergistic opportunities in boundary practices related to the state test, it did not find these opportunities for synergy to rest solely with teachers in classroom work. Rather, the synergistic opportunities were related to several types of actors, including school leaders and curriculum specialists. Consistent with a broadening of instructional responsibility, this study suggests these PD events could be targeted to collaborative teams responsible for student success. An alternative, then, to specific teacher-oriented PD (ex: lessons on developing high-quality assessments and using the state tests) is for additional teachers and curriculum specialists to be included in the boundary practices of developing test items for the state government or some similar forum to those discussed in Chapter 8. Another way to use existing systemic structures to develop local practitioner connections, both at the state level and within local schools and districts, is within the process of reviewing test results in local schools. The development of PD supports for practitioners' review and interpretation of their state test results could be a vehicle for improving their collective understanding of and communication about science education -- a scenario potentially more advantageous than general PD events customized to schools and students.

Another lesson for practice is in the potential use of emerging infrastructures for assessment that are being developed across the state. These efforts seem to indicate greater capabilities for assessment systems are in the near future. However, the extent to which these new information architectures are being designed to support science education was not clear. Across much of this study, the theme of science being subordinated to math and literacy was consistent, suggesting that these new information systems for distributing assessment information may not initially account for any specific characteristics of science education. The challenges for including any specific requirements of science education into these architectures may be exacerbated by the separation of science educators from those involved with developing these architectures, as evidenced by the separation of professional associations. The model in this study highlights both these current limitations and challenges and the potential advantages of their reconciliation.

Contributions to the Literature

As a study that crosses disciplinary boundaries, this study has the potential to inform several literatures. I focus here on four contributions: first, to the literature reviewed in Chapter 2 regarding information use in education; second, to the literature on general theories about educational organizations; third, to theories of organizational synergy and interconnections that were reviewed in Chapter 3; and finally, as a new kind of research method.

The Literature on Information Use in Education:

The literature reviewed in Chapter 2 showed many variations in research location (systemic aperture), grade range, subject area (mostly literacy and math), and perspective (organizational, artifactual, and focused on specific problems/programs). It was also a literature that primarily used case studies arranged in parallel groups, often in hierarchical (nested) constructions, to show contrasts. Furthermore, while these studies looked at the practices of information use, with a few exceptions the artifacts that collect, manage and share that information were not discussed in detail. There was also a methodological tension in this literature, in that those studies looking at wider views of the system tended to report less specific methods.

In addition to contributing to the literature some needed description related to science education and middle grades -- two underrepresented areas -- there are two specific kinds of contributions this model makes.

The first contribution is a flexible research approach that is adaptive to less hierarchical arrangements of research locations. This methodological contribution adds to the growing literature that describes the educational enterprise as being influenced by multiple pathways, different types of agencies, and different types of organizations. These organizational paths can include technical curriculum support structures (Rowan, 1990), coaching (Schon, 1987), professional communities (Lave & Wenger, 1991; Rogers, 1995; Spillane, 2006), commercial organizations (Burch, 2006), and cultural systems (Lee, 1995; Moll, 1998). All of these paths complement the hierarchical (nested) formal structure found in schools, districts, and states.

Most of the research models used in the studies I reviewed can be called *cellular* because of their emphasis on the similar characteristics of the organizational entities, rather than on particular differences, and *cellular* because, other than nesting one element inside another (ex: teachers within schools or schools within districts), they provide little account for the often complexly-interrelated nature of educational systems. This study, in contrast, provides examples of educational practice where there are structures between the school and the state and highlights several of the network structures (ex: professional associations, service agencies, publishers) that can link practitioners.

A second contribution that this study makes to the literature is an emphasis on dynamic qualities. With a few exceptions by Coburn (2004; with Talbert, 2006) and Falk and Drayton (2004), the current literature is a literature *of the moment* that provides little description of developmental trajectories. The model this study produced shows the system in motion. While unable to predict the future, the study uses a theoretical model that incorporates the relevance of developments in technology and diffusion of innovations (Rogers, 1995).

Theories about Educational Organizations: Organic and Dynamic

This research model also contributes to literatures that consider the organizational character of education systems. Weick's (1976) loose-coupling model is an example of a type of general characterization that has been revisited with more specificity in recent years (see Coburn, 2005a; Spillane & Burch, 2004). Rowan's (1990) *Commitment and Control* discussion similarly highlights general characteristics of schooling. By showing heterogeneous organizational structures and relationships, this study presents a more challenging portrait of educational practice. It shows individuals connecting through particular organizational structures (ex: professional associations and state test development processes). These structures show some of the variations that can exist in educational organizations that may not fit so neatly into the types of dichotomous or parallel categories which have become common for theorizing about educational systemic structure.

This study's structure specifically addresses models that use timescale as an organizing principle. In accord with Lemke (2000), and as similarly represented by others (see Erickson, 2007; Hickey & Anderson, 2007), this study presents opportunities for elaboration. Looking at classroom assessment as a low-level ecosocial activity, the review of classroom practices in Chapter 5 and the survey of science teachers discussed in Chapter 6 do show the more frequent temporal patterns at lower ecosocial levels than at higher levels originally described by Lemke (2000). However, they showed that each teacher had a distinct temporal pattern or signature. Three characteristics of this signature were the temporal interval, the regularity of intervals, and the flexibility of the approach. The finding that temporal structure varied across classrooms mirrored what was found across the state government, as well. The annual accountability test was developed according to a two-year development cycle. Looking laterally across the governmental operations, other parallel processes -- for the development of state standards and the development of assessment systems for students with severe disabilities -- had similar long timescales for their work. However, these units were not synchronized to the work of the test development process, providing further support for an ecosocial timescales model with the elaboration of distinct lateral chronic patterns.

The analysis of information platforms in Chapter 9 utilized a diffusion theory framework (Rogers, 1995) that has been applied to a wide range of consumer and organizational scenarios. The organizational landscape discussed in Chapters 7 and 9 presents new opportunities for understanding the diffusion processes, and through them, how different educational organizations behave. Rogers anticipated these opportunities (1995):

> Although it is an important diffusion research tradition in terms of the number of studies completed, education is less important in terms of its contribution to the theoretical understanding of the diffusion of innovations. An exciting potential contribution could be made by the educational research tradition, stemming from the fact that organizations are involved, in one or another, in the adoption of educational innovations. (p. 63.)

This diffusion perspective is one way of understanding educational practice in terms of its dynamic characteristics. Rather than analyzing how *educational systems are* to find ahistorical universals, this perspective asks *how educational systems are evolving*. This dynamic perspective may have broad implications for theories of educational organizations that have often constructed educational practice in constant, rather than developmental, terms.

Contribution to Theories of Systemic Connection

Earlier, in Chapter 3, I discussed different theoretical approaches to organizational action, including the use of artifacts known as boundary objects (Star and Griesemer, 1989; Wenger, 2000), individuals known as brokers (Wenger, 2000) or boundary spanners (Honig, 2003; Tushman & Scanlan, 1981), and boundary practices (Wenger, 2000). All of these theories are concerned with how collaboration and group work occurs. This model provides rich opportunities to view these concepts in practice, where their theoretical perimeters can be stressed. Boundary objects, for example, were originally described in the context of collaboration, where the administrator of a central repository of a natural history museum developed data collection systems to enlist the support of different types of amateurs and professionals whose work was overlapping the work of the museum. Through the development of artifacts for data management (long before the age of electronic documents), the administrator was able to develop cooperative information systems that served multiple interests. At first glance, the assessment item discussed in Chapter 8 could be viewed as a type of boundary object. However, it presents a number of important issues to consider. First, the systems the assessment item operated under were not cooperative, but were imposed by the government. Second, as an artifact, the item is embedded within systems that are part of the policy apparatus and disconnected from the work in schools. Third, as Chapter 9 illustrated, when the textual environment changes, as it did with the data warehouse system used by Hardy, the role of the item appeared to also change. Similarly, the issues of boundary spanners and practices are represented here with more detail and texture than the initial theories provided.

Contributions as a Research Method

As a research method, the network model case study has the ability to be flexible, to develop evolutionary depictions, and to be structured to reach across a large systemic aperture. In situations where research goals call for studying in systemic terms, across related contexts, the network model case-study design is a tool that is a research method designed to adapt to research locations that occur at different scales. The flexible case design in this method allows for research models to support comparisons across contexts and case types. It also benefits from being derived from a well-established methodological family that can be used inside and outside of education, so that it can develop comparisons between educational practice and other types of domains.

Concluding Thoughts: Historical Context, Questions, and a Research Program

I will conclude this dissertation with a reflection on this research from a historical perspective. In important ways, this research is situated within the time in which it was conducted. While most, if not all, professional and scientific research is so situated, occurring within communities that help shape what questions are relevant and what evidence is appropriate at different points in time (Kuhn, 1970), this investigation relates to contemporary developments that mark historical discontinuities in educational practice. These discontinuities cannot easily be held constant for this study. The standards-based reform movement and its latest signature federal NCLB legislation, combined with the emergence of new capabilities from information technology for the collection and distribution of information about students, shape the cusp upon which this research sits.

While historical change often appears in educational research, it is frequently either in the background or within a historical discourse, rather than being considered as a central component in the study of contemporary conditions. For example, when a student in a classroom is observed explaining her understanding of the natural world (McNeill & Krajcik, in press), or a child is studied as he tries to learn from prose and diagrams (Lemke, 1998), living and dynamic communication and linguistic systems are employed. While these systems are evolving, the change is gradual. Similarly, when new configurations of students and teachers are employed, such as to create a community of learners (Brown and Campione, 1994) or to support students jointly scaffolding each other's learning (Erickson, 1996), the study is likely to focus on that unique type of classroom context rather than also considering how classrooms generally have evolved over some period of time. Because of the nature of the relationship between this topic of assessment information to current policy and technological change, this study includes a central component concerned with understanding that change.

At the end of Chapter 2, I discussed the possible emergence of a new field of inquiry in educational research. I based this suggestion on the studies I had reviewed that were raising common issues around the use of data in educational practices, even though they were often looking at different parts of the educational enterprise. These issues included project leadership, infrastructure, organizational capacity (ex: information literacy), and information visualization. I also supported this view by looking at their citations, which drew broadly from what have been historically different fields of research, such as administration and leadership, policy, teaching, and cognition, and from across this small and growing body of literature looking into data use in education. A complementary view to the emergence of a new field is that these studies are indications of a broader movement that is affecting many parts of educational research and society. These changes may be driven by social processes larger than the development of a new type of organizational practice or a professional field to research it. While the proliferation of information communications technologies for student assessment and information within Michigan discussed in Chapter 9 is just one component of this study, there were parallel advances in organizational information communications technology occurring broadly across society during the several decades prior to this study (Friedman, 1995). Could the advances in information infrastructures with the power to disseminate large amounts of information be related to policy expectations (ex: NCLB) in a generative way? That is, could bigger tides of cultural expectations for information availability and instrumental data use be helping to fuel the dissemination of student assessment systems in Michigan that this study discovered?

Broadening the Historical View

Broadening the view beyond Michigan and beyond education, a case could be made that this study was conducted within a period of wide-ranging changes in terms of information and standardization. Michigan's educational sector could be seen as a particular example of evolution, in terms of both capacity and expectations for information, that marks this period.

It was not so many years prior to this study that other advances in communications and network technology, in the form of the World Wide Web/Internet, were often linked to a transformation of many sectors of society (including education) through their power to *visually* present information and connect people across space and time. As Friedman (2005) has noted in *The World is Flat: A Brief History of the Twenty-first Century*, the often *invisible* information technology that links organizations has been responsible for dramatically reshaping both inter and intra-organizational practices around the world. The emergence of standards for data exchange could be another substantial paradigm shift that will be remembered years from now, in addition to the visible effects of browsers, online video, personal representation, and virtual communities.

How can this conception of broad sociological change be deployed in understanding and locating this study historically? Another great communication revolution from an earlier era, the development of the printing press, provides some parallels. Looking back, that revolution was actually comprised of many complex transitions that occurred throughout many aspects of society, rather than being a single dramatic shift. In Eisenstein's (1979) landmark study, she said. "The shift from script to print entailed a large ensemble of changes, each of which are too complicated to be encapsulated by any single formula" (p. 70.) The web of information technology standards, frameworks, and regulations could be a similar ensemble of changes that will increasingly regulate teacher work (Apple & Jungck, 1999) and also may be implicated in the development of less hierarchical educational systems. In ways similar to how the introduction of the printing press to the West reshaped and standardized some established professions (Eisenstein, 1979), systemic reform has been concomitant with the rise of charter schools, home schools, and online schools that have begun to compete with the public system across many parts of the nation. While specifications of performance (ex: learning standards) are intended to create similar minimum outcomes for all students, specific schooling experiences can vary significantly by student. Issues of information use in decision processes abound across these developments. Perhaps the use of information in educational decisions is but one dimension of complexes of transitions occurring more broadly in education and society. It is noteworthy that while most participants in this study, and much of the literature reviewed in Chapter 2, implicated NCLB in the emphasis on information use in school decisions, some of the literature reviewed actually predated NCLB by several years.

One example of a historical realignment that appears within the literature I reviewed is a redistribution of instructional responsibility away from individual teachers towards a collective that includes administrators and collaborative groups (Fink & Resnick, 2001; Leithwood & Riehl, 2003; Little, 2007). A recent body of work at The Center for the Study of Teaching and Policy that includes Copland and Knapp's (2006) *Connecting Leadership with Learning: A Framework for Reflection, Planning, and Action* explicated this reshaping of educational organizations and the role that information at all levels plays in this new way of viewing educational practice. Introducing their comprehensive approach, they describe the shift towards universal achievement that is so much a part of the current reform era:

> Creating an agenda for equity and excellence in student achievement is of the utmost importance in the current accountability landscape. Despite the urgency embedded in this statement, such a *shift in thinking* presents new challenges that were previously unknown or perhaps not even considered for public education. Our schools need leaders who can *systematically* deliver on the promise of instructional expertise and high-quality learning for all children. To be successful, school leaders will need new *more powerful tools* to support their efforts to develop the most effective and sustainable leadership strategies. (Emphasis added, p. x.)

Throughout their work, especially in two recent publications, *Data-informed Leadership in Education* (Knapp, Swinnerton, & Monpas-Huber, 2006) and *Data-informed Leadership: Insights from Current Research, Theory, and Practice* (Knapp, Copland, and Swinnerton, 2007), these authors describe the use of information in decision-making in ways consistent with the work reviewed in Chapter 2. They also describe the relationship between practices and artifacts in historical terms, noting the previous several decades, the same period when information technology became more widely available, as significant:

> . . . the data dialogue has entered a new era in which leaders' engagement in data-based problem-solving is benefiting from new tools and trends not previously known. Building on a robust evaluation movement in the 1960s and 1970s, a variety of techniques and strategies are now available for systematically evaluating the implementation, effects, and effectiveness of educational programs, policies, or initiatives…. the rapid sophistication of technologies for handling digital information make the prospects for educational decisions rooted in relevant evidence more realistic, yet simultaneously more costly and complex. (p. 74.)

Indeed, the authors present the development of accountability policies, evaluation methods, and the technical tools for supporting these practices as *parallel processes*, which could help explain the emergence of both leadership and technology issues across the literature reviewed earlier, as both are key elements of related complexes of social processes that mark this reform period.

Questions for a More Infrastructure-driven Education Climate

What might this period of accountability mean in terms of issues and problems to be considered? I approach this question by recalling some changes related to technology evident in this study. In addition to the technologies being implemented at the district level, the ones that supported systemic uses of assessment information, there were also cases of technology innovation by teachers using course websites, PowerPoint, and various assessment-related classroom tools to support their instruction. The study's model then documented a time where electronic textual capabilities were continuing to proliferate with no end in sight. This case study comprised of related case studies, from classrooms to the state test, is in important ways a snapshot of educational components undergoing technological evolution that has implications for systematicity. Systematicity, however, can be used in a number of ways: as artifacts, as individuals and organizations connected by artifacts, and as reciprocal action. Across these three uses of the term "system" are three different types of questions we can use to consider these implications. These innovations relate to the conceptions of systematicity in different ways.

From an artifactual perspective, the diffusion of assessment infrastructures in administrative levels, combined with the influence of textbook manufacturers described in Chapter 9, indicates a possible future scenario where across the state, districts could be utilizing a small number of student information infrastructures. This scenario would parallel other software niches, such as for payroll and accounting, that at one time had many solutions but over time became markets dominated by a small handful of large vendors offering very similar products. Within this view, in addition to corporate efficiencies from smaller numbers of software products, we can see the individual practitioners as potentially subjected to increasing levels of homogenization and control by state-centered systemic apparatuses. As discussed in Chapter 7, the state was a partner in this process by developing standards for information exchange. Since the state is in a position of helping to constitute these infrastructures, does the state's role represent areas for concern regarding systemic control?

The artifactual view may lead towards questions about the agentive and possible political role of these information systems. Winner's (1986) essay *Do Artifacts Have Politics* described how a manager of public works projects in New York City specified that bridges being built over an expressway leading to beaches and public attractions be made too low for buses to pass through. The result of this was that people of low income were excluded from certain recreational opportunities by virtue of the bridges' designed height. While this case is extreme, it does point out that infrastructures that seem innocuous and everyday may have dimensions that affect different populations in far different ways. In a period where commercial interests provide resources for instruction that are aligned with systems of measurement they also provide, how well will the public's educational interests be served?

It is possible to envision that the practitioners using these technologies that Bowker and Star (1999) refer to as *boundary infrastructure* might find their practice shaped by the infrastructures' technical designs in terms of the codes, categories, and procedures embedded within the technology. Lessig, in *CODE and Other Laws of Cyberspace* (1999) argues, that software architecture can both enable and restrict through its often unseen design. This is Winner's examples in digital form. While this is an extreme position, and not one strongly supported in the nascent research on educational practice reviewed in Chapter 2, it is a possibility that might raise concerns for practitioners, especially science educators, who seem to have had little involvement in the infrastructural process during the time of the study.

As the focus turns to the third use of the term "system," involving reciprocal action -- arguably the most important one from a policy perspective -- an important value question emerges: is a system for education where policies and infrastructures support consistent practice desirable? The answer any person comes up with will likely be influenced by his or her world view. Those who believe in education as a fundamentally manageable process, who look for greater systemic control and greater alignment with standards and see systems that deliver information about student performance as essential mechanisms for accountability, will be in favor. Those who would tend to believe in education as a process that requires and takes advantage of individual human differences and innovation may hold different views. From this perspective, a homogenizing complex of systems backed by policy that contains instrumental force may not be so good. From this perspective, a system that implements uniformity and coherence could have significant disadvantages, as argued by Moss (2004). This perspective may bring American education into contact with its history (and in some cases its present) marked by industrial epistemologies that classified students by inherent traits and tracked them into instructional groups according to systemically-defined potential (Lagemann, 2000).

Extending the Research: Synergy, Scaling, Description, and Design

There are many ways this research can be extended. The selection of teachers, schools, and states could be expanded, and then the types of claims that could be made about educational practice could be strengthened. More observations, specifically of boundary practices and instruction, would allow a sharper depiction of both the key synergistic processes and also the sites where synergistic results need to appear. Further, researching over longer periods of time and with more practitioners of different types would help expand the depth of model produced. Studying over a longer time period, and looking specifically at organizations as they move from a period of low information literacy through the integration of information into organizational routines, would provide an important lens on adoption and resistance -- two key systemic issues. In conjunction with an up-close study of information literacy practices, study of the visualization and representation of data could make an important contribution to an area that has thus far been only slightly addressed in educational research, and little outside of it.

This study has been descriptive. Another possible direction to take this research is in the area of design. Design-based research can be seen as an educational research branch of the design sciences (Simon, 1996). Design studies have been used in recent years to understand how innovations in classroom practice can improve student learning (Collins, Joseph, & Bielaczyk, 2004). Design-based research involves studying an innovation inserted into a social context or activity structure. The research begun in this dissertation may benefit from translating these principles of design-based research into organizational contexts so as to understand how different educational actors may respond to different configurations of standards and presentations of assessment information. The area of assessment information use, with its wealth of visualization options and different topics, would perhaps benefit from deliberate design research that can test alternative information architectures in practice. This research would develop research-informed information systems and device interpretation activities to improve practitioner synergy around issues of science education.

From the perspective of individuals, this study also showed evidence of pressures for uniformity. Not only was the emphasis on state content standards from NCLB evident, but also the trend for teaching practice to be less private and more open to inspection, control, and performance standards through the broad realignment of instructional responsibility was part of the model. In thinking of collaborative processes and the code-driven information systems that support them, one area that calls for study is how groups of educational professionals may make their decisions based on information presented in complex visual assemblages of the kind shown in Chapter 9. The professional vision that these displays support is important to understand from the perspectives of different types of systemic actors.

For both the systemic and the design perspectives, one characteristic of this study that will likely be present in any continuation of this research is its attention to change. Whether adopting a purely descriptive or a design-based stance, this type of research, which looks at education in large units from a systemic frame and where information of any kind is an important part of the systemic construction, will likely be research that is contending with a dynamic policy climate and rapidly evolving technologies, rather than with the kind of stable research paradigms and small numbers of research locations found in some other aspects of educational research.

Closing Questions

A question remains as to where the continuation of this research would belong in disciplinary terms. The literature I reviewed in the beginning of this dissertation may either index the emergence of a new field or index changes in existing fields. While the investigations reviewed as part of this study could form the basis for an emerging field that this study could join, an alternative is possible. The breadth of disciplinary representation – from policy to cognition to technology – in the articles reviewed could indicate de-centering forces that may pull investigations into the use of information for educational decisions back toward more traditional domains and away from a cohesive research discourse around these systemic issues. It is also conceivable that these complex social processes may eventually spawn discourses about data and information in education in historically established domains. Then, the research regarding information use for literacy and mathematics, for example, might be represented within those types of forums or in leadership and policy discussions, rather than in forums dedicated to understanding how artifacts and educational organization interact across domains.

Regardless of how this type of model-based research is extended, its fundamental strength and challenge will be in issues of systematicity and scale. However extended, when looking at issues of systematic functioning, the models produced from research along these lines will almost certainly encounter the often dichotomous tension between the top-down and bottom-up views. Whether seen as an engine that runs without producing desired effects or as a machine with the potential to restrict diversity, a model of educational systems may play into political tensions. Furthermore, it is evident from this research that, regardless of the disciplinary home for a systemic model, systemic issues can cross terrains of geography, role, and organization. Likewise, systematicity can be found in the relations between the human, social, and organizational realms and between technologies and texts. Research that looks at both production and consumption of assessment information, as this one does, must overcome challenges of scale and aperture in attempting to follow increasingly digitalized texts across contexts and over time.

Appendix A – Master Participant List

Below is a list of the study participants and the amount of time they were studied and the number of cases they participate in as primary participants or supplemental informants. The Code is the abbreviation used in transcripts for their talk. The members of the three committees were represented in transcripts with a number (ex:R2 for reviewer 2).

Code	Pseudonym	Pseudonym	Cases	Type	Hours
AD	Angela DuBous	Hardy Middle School	1	Informant	2.00
AG	Art Gnatt	Michigan	1	Informant	0.58
AP	Ann Pobzerniac	Hardy Middle School	1	Informant	3.00
AT	Angie Toliver	Galenan ISD.M/S Center	1	Informant	0.17
AV	Annie Vanduesen	Green Glenn Pub Schools	3	Informant	0.33
BE	Bob EnSpania	Avon Falls Middle School	2	Primary	1.25
BIAS	Bias committee (7)	Michigan	1	Informant	1.00
BL	Bernie Lauer	Mapleland ISD	3	Primary	2.42
BR	Ben Raminskis	Crimson Middle School	2	Primary	1.00
BS	Bob Senoff	Hardy Middle School	5	Primary	4.67
BT	Betsy Dearing	Avon Falls Middle School	3	Primary	7.08
BW	Burt Wainwright	Swallow Middle School	3	Primary	6.58
CA	Cathy Amazingly	Swallow Middle School	3	Informant	4.00
CAC	Content committee (8)	Michigan	0	Informant	1.00
CC	Christy Connolly	Crimson Middle School	4	Primary	2.18

Code	Pseudonym	Pseudonym	Cases	Type	Hours
CD	Carmen DuRonder	Hardy Middle School	2	Informant	0.50
CE	Corrine Eaton	Challenge Run M/S Center	3	Primary	1.20
DD	Donna Dinard	Dorchester Middle School	2	Informant	0.92
DH	Debbie Huston	Hardy Middle School	0	Informant	0.25
DJ	Don Justice	Michigan	1	Informant	0.58
DP	Don Pulte	Tichiochi ISD	2	Primary	0.45
DS	Duane Sprocket	Hardy Middle School	2	Informant	2.00
DV	Diana VanderMiller	Hardy Middle School	3	Primary	5.00
EB	Ed Bedminster	Crimson Middle School	2	Informant	1.38
FC	Faith Churchill	Dorchester Middle School	2	Primary	1.67
FG	Focus Group (6)	Michigan	2	Informant	2
GB	George Brunson	Avon Falls Middle School	1	Informant	2.83
JB	Jane Bailey	Swallow Middle School	1	Informant	0.15
JF	Janey Fess	Avon Falls Middle School	2	Primary	3.63
JH	Jim Heinrich	Avon Falls Middle School	2	Primary	4.08
JM	Joshua Martinique	Michigan	4	Primary	3.17
JT	Jim Rontonovich	MAISA	1	Informant	0.17
KM	Karen Minor	Tichiochi ISD	1	Primary	0.33
KR	Keith Roberts	Michigan	1	Informant	1.00
MW	Mike Wolson	SuccessLine	3	Informant	3.33
NN	Nancy Newman	Wushutunesa ISD	8	Primary	5.00
PB	Paul Bond	Swallow Middle School	2	Primary	2.75
PD	Pete Darmond	Science Teacher Community	5	Primary	1.08
PH	Patty Hollander	Michigan	0	Informant	0.58
POLIT	7members	Michigan	0	Informant	

Code	Pseudonym	Pseudonym	Cases	Type	Hours
PS	Pete Stomber	Michigan	0	Informant	2.00
RB	Bob Black	Michigan	1	Informant	4.00
RE	Roscoe Ells	Michigan	2	Primary	3.75
SC	State committee (7)	Michigan	0	Informant	3.00
SD	Stan Dubovski	Avon Falls Middle School	3	Primary	2.50
SH	Sally Henrig	Hamilton ISD	0	Informant	0.25
TD	Toni Donard	Swallow Middle School	1	Informant	4.00
VD	Vance Dorn	Michigan	0	Informant	1.00
VJ	Valerie Jones	Hardy Middle School	1	Primary	2.00

Appendix B – Master Case Structure List

Below is a list of the case structure that shows which participants contributed to which cases as primary participants and as informants.

Case name	Case description	Case Relationship	Participant Name	Official Role
Individuals				
Bernie Lauer	Assessment and evaluation specialist			
		Primary	Bernie Lauer	Other
		Informant	Nancy Newman	Curriculum
Ben Raminskis	Classroom teacher			
		Primary	Ben Raminskis	Teacher
Betsey Dearing	Classroom teacher			
		Primary	Betsy Dearing	Teacher
Bob EnSpania	District Superintendent			
		Primary	Bob EnSpania	Executive
Bob Senoff	Head of assessment and evaluation			
		Primary	Bob Senoff	Executive
Burt Wainwright	Principal and curriculum coordinator			
		Primary	Burt Wainwright	Principal
		Informant	Cathy Amazingly	Teacher
Christy Connolly	Classroom teacher			
		Primary	Christy Connolly	Teacher
		Informant	Ed Bedminster	Principal
		Informant	Nancy Newman	Curriculum
		Informant	Annie Vanduesen	Other
Corrine Eaton	Science Education			

Case name	Case description	Case Relationship	Participant Name	Official Role
	Leader			
		Primary	Corrine Eaton	Curriculum
		Informant	Pete Darmond	Curriculum
Dianne Vandermiller	Coach and assistant principal			
		Primary	Diana VanderMiller	Curriculum
		Informant	Bob Senoff	Executive
		Informant	Carmen DuRonder	Principal
Don Pulte	Assistant Superintendent			
		Primary	Don Pulte	Executive
		Informant	Nancy Newman	Curriculum
Faith Churchill	Classroom teacher			
		Primary	Faith Churchill	Teacher
		Informant	Donna Dinard	Principal
Hardy Kit Teacher	The classroom teaching practices of the kit teacher at Hardy MS			
		Primary	Valerie Jones	Teacher
		Informant	Bob Senoff	Executive
		Informant	Diana VanderMiller	Curriculum
		Informant	Duane Sprocket	Teacher
		Informant	Annie Vanduesen	Other
Janey Fess	Curriculum Director			
		Primary	Janey Fess	Curriculum
Jim Heinrich	Classroom teacher			
		Primary	Jim Heinrich	Teacher
Joshua Martinique	Manager of state assessment and former ISD consultant			
			Joshua Martinique	State
Karen Minor	Curriculum Director			
		Primary	Karen Minor	Curriculum
		Informant	Nancy Newman	Curriculum

Case name	Case description	Case Relationship	Participant Name	Official Role
Nancy Newman	Curriculum Director			
		Primary	Nancy Newman	Curriculum
Paul Bond	Classroom teacher			
		Primary	Paul Bond	Teacher
		Informant	Cathy Amazingly	Teacher
		Informant	Angie Toliver	Curriculum
Pete Darmond	Science Education Leader			
		Primary	Pete Darmond	Curriculum
Roscoe Ellis	MEAP Science Consultant			
		Primary	Roscoe Ells	State
		Informant	Bob Black	State
		Informant	Burt Wainwright	Principal
		Informant	Bob Senoff	Executive
Stan Dubovski	Principal at Swallow Middle School			
		Primary	Stan Dubovski	Principal
Schools				
Avon Falls	Middle School			
		Primary	Stan Dubovski	Principal
		Informant	Bob EnSpania	Executive
		Informant	Betsy Dearing	Teacher
		Informant	Jim Heinrich	Teacher
		Informant	George Brunson	Teacher
		Informant	Janey Fess	Curriculum
Crimson	Suburban Middle School			
		Primary	Ed Bedminster	Principal
		Informant	Annie Vanduesen	Other
		Informant	Ben Raminskis	Teacher
		Informant	Christy Connolly	Teacher
		Informant	Nancy Newman	Curriculum
Dorchester	Affluent Rural School			
		Primary	Donna Dinard	Principal
		Informant	Faith Churchill	Teacher
Hardy	Large Urban Middle School			
		Primary	Diana	Curriculum

Case name	Case description	Case Relationship	Participant Name	Official Role
		Informant	VanderMiller Ann Pobzerniac	Teacher
		Informant	Bob Senoff	Executive
		Informant	Duane Sprocket	Teacher
		Informant	Carmen DuRonder	Principal
		Informant	Mike Wolson	Vendor
		Informant	Angela DuBous	Teacher
Swallow	Small Middle School			
		Primary	Burt Wainwright	Principal
		Informant	Toni Donard	Teacher
		Informant	Cathy Amazingly	Teacher
		Informant	Jane Bailey	Principal
		Informant	Paul Bond	Teacher
Professional Communities				
Assessment Community	Community of state test producers/consumers			
		Primary	Bernie Lauer	Other
		Informant	Mike Wolson	Vendor
		Informant	Nancy Newman	Curriculum
		Informant	Joshua Martinique	State
School Administration Community	Association of School Administrators (combines MAISA and MASA)			
		Primary	Jim Rontonovich	Executive
		Informant	Bernie Lauer	Other
		Informant	Joshua Martinique	State
		Informant	Don Pulte	Executive
		Informant	Nancy Newman	Curriculum
Science Teachers Community	Michigan Science Teachers Association			
		Primary	Pete Darmond	Curriculum
		Informant	Corrine Eaton	Curriculum
		Informant	Focus Group	Focus Group

Case name	Case description	Case Relationship	Participant Name	Official Role
		Informant	Betsy Dearing	Teacher
		Informant	Christy Connolly	Teacher
State and Other				
MEAP Development	State test development process			
		Primary	Roscoe Ells	State
		Informant	Joshua Martinique	State
		Informant	Christy Connolly	Teacher
		Informant	Focus Group	Focus Group
		Informant	Corrine Eaton	Curriculum
		Informant	Pete Darmond	Curriculum
		Informant	Bias committee	Committee
		Informant	Don Justice	Other
		Informant	Keith Roberts	State
Successline	Software Vendor			
		Primary	Mike Wolson	Vendor Rep
		Informant	Pete Darmond	Curriculum
		Informant	Stan Dubovski	Principal

Appendix C – Master Event List

Below is a list of the study data collection events and the participants at them. Not listed are email contacts and web-based research that did not involve participants.

April 2005

4/15/2005 Initial meeting 20.00 Min.

Participants none

Bob EnSpania

December 2006

12/19/2006 Inventory Interview 55.00 Min.

Participants Fieldnotes

Burt Wainwright

12/20/2006 Planning meeting 30.00 Min.

Participants Fieldnotes

Bob EnSpania

Janey Fess

Stan Dubovski

12/19/2006 Initial meeting 45.00 Min.

Participants Fieldnotes

Paul Bond

January 2007

1/16/2007 Recorded Interviews Dirk 100.00 Min.

Participants Video

Burt Wainwright

1/17/2007 Initial Meeting with Curriculum Person 60.00 Min.

Participants Fieldnotes

Janey Fess

1/17/2007 Initial meeting with OEAA 120.00 Min.

Participants Fieldnotes

Joshua Martinique

Pete Stomber

Roscoe Ells

1/18/2007 Initial Interviews with 45.00 Min.
Participants Fieldnotes
Stan Dubovski
1/29/2007 MEAP Item 60.00 Min.
Bias/Sensitivity Review
Participants Video
7 members bias committee
Bob Senoff
Keith Roberts
Rob Edbar
1/30/2007 Martin science 0.00 Min.
department meeting
Participants None
Burt Wainwright
Cathy Amazingly
Jane Bailey
Paul Bond
Toni Donard
1/30/2007 MEAP Content Review Committee 60.00 Min.
Participants Video
8 member content committee
Roscoe Ells
1/30/2007 Interviews with Tom 60.00 Min.
Participants Fieldnotes
Jim Heinrich
1/31/2007 Video interview Stan 45.00 Min.
Participants Video
Stan Dubovski
1/15/2007 Meet with Vince Dean 30.00 Min.
Participants Fieldnotes
Vance Dorn
1/18/2007 Initial Interview 45.00 Min.
Participants Fieldnotes
Jim Heinrich
1/18/2007 Initial Interview with 45.00 Min.
Betsy
Participants Fieldnotes
Betsy Dearing
1/18/2007 Interview with Bill 25.00 Min.
Participants Video
Bob EnSpania
1/29/2007 Interview 45.00 Min.
Participants Video

	Roscoe Ells		
1/30/2007	Interview in room		90.00 Min.
	Participants	Video	
	Betsy Dearing		
1/31/2007	Video Interview with Betsy		120.00 Min.
	Participants	Video	
	Betsy Dearing		
1/31/2007	Interview with Janey Fess		106.00 Min.
	Participants	Video	
	Janey Fess		
February 2007			
2/7/2007	Martin science meeting makeup		120.00 Min.
	Participants	Video	
	Burt Wainwright		
	Cathy Amazingly		
	Paul Bond		
	Toni Donard		
2/12/2007	Phone conversation Judy Foss		22.00 Min.
	Participants	Audio	
	Janey Fess		
2/15/2007	Meet to discuss MEAP item review process		60.00 Min.
	Participants	Video	
	Bob Black		
2/15/2007	Meet Connie Duncan BCMSC		52.00 Min.
	Participants	Audio	
	Corrine Eaton		
2/18/2007	Meet to discuss Special Ed Process		30.00 Min.
	Participants	Audio	
	Vance Dorn		
2/16/2007	K-12 Meeting Lunch		30.00 Min.
	Participants	None	
	Betsy Dearing		
	George Brunson		
	Jim Heinrich		
	Stan Dubovski		
2/9/2007	Discuss Calhoun ISD's Data Warehouse		15.00 Min.
	Participants	Fieldnotes	

Sally Henrig

2/8/2007	Chat on phone		20.00 Min.
	Participants	Audio	
	Corrine Eaton		
2/27/2007	MSTC Data Warehouse Presentation		60.00 Min.
	Participants	Audio	
	Bernie Lauer		
2/27/2007	Lunch at MSTC		40.00 Min.
	Participants	Fieldnotes	
	Bernie Lauer		
2/27/2007	Phone conversation		45.00 Min.
	Participants	Audio	
	Bernie Lauer		

March 2007

3/13/2007	K-12 Meeting		120.00 Min.
	Participants	Audio	
	Burt Wainwright		
	Cathy Amazingly		
	Toni Donard		
3/6/2007	Observe Tracker training		180.00 Min.
	Participants	Video	
	Ann Pobzerniac		
	Bob Senoff		
	Diana VanderMiller		
	Mike Wolson		
3/14/2007	Initial Interview		90.00 Min.
	Participants	Video	
	Diana VanderMiller		
3/21/2007	Interview		30.00 Min.
	Participants	Audio	
	Carmen DuRonder		
	Diana VanderMiller		
3/14/2007	Interview with Bill		40.00 Min.
	Participants	Audio	
	Bob Senoff		
3/21/2007	Teacher discussion		120.00 Min.
	Participants	Audio	
	Angela DuBous		
	Duane Sprocket		
	Valerie Jones		
3/19/2007	Initial meeting		45.00 Min.
	Participants	Video	

Christy Connolly

3/15/2007 Initial meetings 60.00 Min.

Participants None

Ben Raminskis

3/22/2007 Interview with Christy 46.00 Min.

Participants Video

Christy Connolly

3/22/2007 Interview with Ed 30.00 Min.

Participants Audio

Ed Bedminster

3/20/2007 Initial Interview 50.00 Min.

Participants Fieldnotes

Faith Churchill

3/23/2007 Recorded interview 50.00 Min.

Participants Audio

Faith Churchill

3/23/2007 Interview Martineau 70.00 Min.

Participants Audio

Joshua Martinique

3/23/2007 Interview with Ed 28.00 Min.

Participants Audio

Ed Bedminster

April 2007

4/4/2007 MDE Stat Review Committee 180.00 Min.

Participants Video

6 member stat committee (7 members)

Bob Black

4/16/2007 Questionnaire with Ed 25.00 Min.

Participants Audio

Ed Bedminster

May 2007

5/21/2007 Initial phone conversation 65.00 Min.

Participants Fieldnotes

Pete Darmond

November 2007

11/15/2007 Followup interview 40.00 Min.

Participants Audio

Christy Connolly

Appendix D – Master Question List

Below is a working list of the questions that were asked in interviews and the Michigan Science Teachers Association (MSTA) survey.

Question	Name *Question Text*
1	Assessments used through calendar *One thing I want to understand is the relationship between assessments and the school year. I have here an illustration of school years showing typical vacations and breaks. Feel free to refer to this illustration and to mark on it as you describe how y*
2	Student timeline *"We will approach this issue of assessments from another perspective now, the student. Using this illustration of a student's path through the grades, can you indicate important points in time for assessment of their knowledge or learning that will help "*
4	Science support structure *Now I would like to get your view of the organizational structures that exist in this school/district for helping middle school science teachers.*
5	Assessment info: inside-inside *"Teachers use a variety of methods to assess their student learning in science. The uses quizzes and test and projects to name a few common examples. What do they do with these assessments, what different purposes do they serve or could they serve?"*
6	Clarification text for Inside-outside question *"I am going to use some terms of inside and outside the classroom in some questions that follow. What I mean by inside the classroom are assessments that are managed as part of the teacher's instructional practice. So, even if a teacher takes an assessment"*
7	Assessment info: inside-inside *"Teachers use a variety of methods to assess their student learning in science. The uses quizzes and test and projects to name a few common examples. What do they do with these assessments, what different purposes do they serve or could they serve?"*
8	Assessment info: outside--inside *"There are also assessments that are not assigned by a teacher. Some of these are state and national tests and we can also view other tests given in a school system as possibly relevant to the work that teachers do ?*
9	Assessment info: inside-inside *"Thinking about the organization, under what circumstances and at what times can you see assessments that teachers develop or assign such as quizzes*
10	Multidimensional space/learning progression *"I have a hypothetical question. Suppose there is a new type of assessment that simultaneously measures student ability in multiple areas rather just giving a score for biology or chemistry. suppose there were three scores"*
11	In the next five years *"In the next five years, what are some of the important changes in your job regarding assessments ?"*
12	How is NCLB impacting your work with science education *There is a lot of discussion now of No Child Left Behind and the impact it is having on the work that goes on in schools. Can you describe how you see NCLB affecting your work and its lasting impact*

on education?

13 How would this decade be described in terms of change

"The last decade or so has seen a lot of major developments in technology and in the approach to schooling. Thinking about information and communication, how have you seen this period affecting your work as it relates to understanding student learning?"

14 New assessment for science education what and where

If you could have some new forms of assessment relevant for science education designed and then made available for you to use at any part of the educational process or system. What would be your top priorities? How would what you want today be different

15 Limitations of assessments

What part of what you care about are assessments both inside and outside the classroom least suited to measure?

16 Benefits of assessments

What part of what you care about are assessments from inside and outside the classroom best suited to measure?

17 Shortest time between measurements

What is the shortest amount of time a student should be allowed to be learning without some kind of easurement?

18 Longest time between measurements

What is the longest amount of time a student should be allowed to be learning without some kind of measurement?

20 Years experience

21 What types of students are in your district?

What types of students are in your district?

22 District size

Approximately how many students are in your district?

23 Roles

What roles do you currently perform (check all that apply) ?

24 Work location

Where do you work? (check all that apply)

25 School type

Your school is a: (check all that apply)

26 School size

Number of students in school

27 Science teachers in school

Number of science teachers in school

28 MSTA members

Number of teachers who are MSTA members

29 Grades responsible for

For which grades are you responsible (check all that apply)?

30 Subjects responsible

For which subjects are you responsible (check all that apply)?

31a Experience as a science teacher

How much experience do you have?

31b Experience as a teacher

How much experience do you have?

31c Experience as a dept head

How much experience do you have?

31d Experience as curriculum specialist

How much experience do you have?

32 MEAP Experience

Have you ever been involved Michigan Department of Education (MDE) committees/events where MEAP items have been developed or reviewed?

33 Reasons for classroom assessments

What are the reasons that teachers typically give students assessment tasks (check all that apply)?

34a Student Performance: Interest

How important is it for students to be able to perform on the following areas in their classroom assessments?

34b Student Performance : Terminology

How important is it for students to be able to perform on the following areas in their classroom assessments?

34c Student Performance: Scientific communication

How important is it for students to be able to perform on the following areas in their classroom assessments?

34d Student Performance: Interpret diagrams

How important is it for students to be able to perform on the following areas in their classroom assessments?

34e Student Performance: Draw diagrams

How important is it for students to be able to perform on the following areas in their classroom assessments?

34f Student Performance: Modeling

How important is it for students to be able to perform on the following areas in their classroom assessments?

35a Taking Classroom Assessments: Multiple opportunities formative

How important is it for students when taking classroom assessments to have: multiple opportunities to succeed on regular (ex: daily or weekly) assessments?

35b Taking Classroom Assessments: Multiple opportunities summative

How important is it for students when taking classroom assessments to have: multiple opportunities to succeed on end of unit or yearly assessments?

35c Taking Classroom Assessments: achievable

How important is it for students when taking classroom assessments to have: some tasks that are achievable?

35d Taking Classroom Assessments: Push limits

How important is it for students when taking classroom assessments to have: tasks that push the limits of what they can do?

36 Assessment differences

A recent article showed that student's grades and their performance on standardized tests were often quite different

37 MEAP Discussions

With whom do you discuss the MEAP results with when they are released (check all that apply)?

38 MEAP Information

How do you get to see the MEAP results (check all that apply)?

39a MEAP: Before year

How much is the MEAP a factor in the work you do at different times of the year?: Before the school year

39b MEAP: Start of year

How much is the MEAP a factor in the work you do at different times of the year?: At the beginning of the year
before the test

39c MEAP: During test

How much is the MEAP a factor in the work you do at different times of the year?: During the test

39d MEAP: After test

"How much is the MEAP a factor in the work you do at different times of the year?: After the tests

39e MEAP: After results

How much is the MEAP a factor in the work you do at different times of the year?: After the results are

39f MEAP: End of year

How much is the MEAP a factor in the work you do at different times of the year?: At the end of the year

40 Other assessments

"In addition to the MEAP and classroom assessments

41a CA frequency: Bellwork

Please take a few moments to tell us about your classroom assignments.

41b CA frequency: Quizzes
Please take a few moments to tell us about your classroom assignments.
41c CA frequency: Worksheets
Please take a few moments to tell us about your classroom assignments.
41d CA frequency: Unit Tests
Please take a few moments to tell us about your classroom assignments.
41e CA frequency: Formal Exams
Please take a few moments to tell us about your classroom assignments.
41f CA frequency: Projects
Please take a few moments to tell us about your classroom assignments.
41g CA frequency: Other
Please take a few moments to tell us about your classroom assignments.
42a Source of CA
The questions I use in my classroom come from (check all that apply):
42b CA Influences: Pre-service
How influential were the following sources in your approach to classroom assessment: Your pre-service training
42c CA Influences: In-service
How influential were the following sources in your approach to classroom assessment : In-service professional development
42d CA Influences: Colleagues
How influential were the following sources in your approach to classroom assessment : Colleagues approaches to assessment
42e CA Influences: First experience
How influential were the following sources in your approach to classroom assessment?: Your first teaching experiences
42f CA Influences: District
How influential were the following sources in your approach to classroom assessment?: District/science center recommendations
42g CA Influences: Textbooks
How influential were the following sources in your approach to classroom assessment?: Textbooks
42h CA Influences: Trial and Error
How influential were the following sources in your approach to classroom assessment?: rial and Error
43 Assessment Revision
"Over the last few years
44 Question Difficulty
"When a student is taking a test
46 MEAP Meetings
How many meetings do you have each year where the MEAP results are the focus?
46 MEAP Relevance
How would you characterize MEAP results in terms of relevance to your teachers?
47 MEAP Confidence
How confident are you that you are able to use the MEAP results to shape instruction?
48 MEAP Routines
How much of what you did last year to use MEAP results was new?
49a MEAP Dev: Writing
How many times have you been involved in the following?: MEAP Item writing
49b MEAP Dev: Bias
How many times have you been involved in the following?: Reviewing MEAP items for bias/sensitivity
49c MEAP Dev: Content
How many times have you been involved in the following?: Review items for content
49d MEAP Dev: Standards
How many times have you been involved in the following?: Development of standards and grade-level expectations
49e MEAP Dev: General
How many times have you been involved in the following?: General MEAP Item Review

50a MEAP Exp Benefits: Process
How useful has your participation in MEAP events been for improving your understanding of?: The MEAP process?

50b MEAP Exp Benefits: School
How useful has your participation in MEAP events been for improving your understanding of?: What happens at your school/job?

50c MEAP Exp Benefits: Others
How useful has your participation in MEAP events been for improving your understanding of?: What occurs at other participants' schools?

50d MEAP Exp Benefits: HQ Assessments
How useful has your participation in MEAP events been for improving your understanding of?: High-quality assessment?

51a MEAP Dev Qual: SMK
How important are these characteristics for participating in MEAP item development and review?: Deep

51b MEAP Dev Qual: STE
How important are these characteristics for participating in MEAP item development and review?: Teaching

51c MEAP Dev Qual: Breadth
How important are these characteristics for participating in MEAP item development and review?: Broad

51d MEAP Dev Qual: Gen Exp
How important are these characteristics for participating in MEAP item development and review?: Broad

51e MEAP Dev Qual: Mindset
How important are these characteristics for participating in MEAP item development and review?: A certain

51f MEAP Dev Qual: Committee
How important are these characteristics for participating in MEAP item development and review?: Good

52 MEAP Expert
Are you considered an expert in the MEAP process by your school because of your involvement in MEAP development?

53 MEAP Comments
Enter any comments about what you think of the process of participating in the MEAP development/review. For whom would you recommend this experience?

54 Gender
What is your gender?

55 Last 5 years
"Over the last five years

56 NCLB
How much is No Child Left Behind (NCLB) affecting your job?

57 MEAP Downside
Please describe what you see as the most harmful aspects of MEAP testing. How does the MEAP make your job more difficult?

58 MEAP Upside
Please describe what you see as the most beneficial aspects of the MEAP program. How does it help you or education?

59 Principal Involvement
How involved is your school principal or administrators in issues related to science assessment currently?

60 Other comments
Please feel free to make any other remarks to the researcher about assessment for science.

19 MSTA Region
Which MSTA Region are you in

102 Which teachers resist assessment/accountability

What types of teachers resist accountability? Are there any patterns?

103 When MEAP results are released to school what occurs?

When MEAP results are released to school what occurs?

114 How important is terminology for science teaching?

How important is terminology for science teaching?

References

AAAS, American Academy for the Advancement of Science. (1993). *Benchmarks for Science Literacy*. New York: Oxford University Press.

Anderson, C. (2004). Science Education Research, Environmental Literacy, and Our Collective Future. *NARST News, 4*.

Apple, M. W., & Jungck, S. (1999). "You Don't Have To Be A Teacher To Teach This Unit": Teaching, Technology and Control in the Classroom. In Bromley, H. & Apple, M. (Eds.), *EDUCATION/TECHNOLOGY/POWER: Educational Computing as a Social Practice* (pp. 133-56). Albany, NY: State University of New York Press.

Airasian, P. (1988). Measurement Driven Instruction: A Closer Look. Educational Measurement: Issues and Practices, 7(4), 6-11.

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: what is -- or might be -- the role of curriculum materials in teacher learning and instructional reform? Educational Researcher, 25(14), 6-8.

Bass, F. (1969). A New Product Growth for Model Consumer Durables. Management Science, 15(5), 215-227.

Bennett, R. E. (2002). Inexorable and Inevitable: The Continuing Story of Technology and Assessment. *Journal of Technology, Learning, and Assessment*, *1*(1).

Bernhardt, V. (1998). Data Analysis for Continuous School Improvement. Larchmont, NY: Eye On Education.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education, 5, 7-74.

Boote, D., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. Educational Researcher, 24(6), 3-15.

Boote, D., & Beile, P. (2006). On "Literature Reviews of, and for, Educational Research": A Response to the Critique by Joseph Maxwell. Educational Researcher, 55(12), 32-35.

Boudett, K., City, E., & Murnane, J. (Eds.). (2005). Data Wise: A Step-by-Step Guide to Using Assessment Results to Improve Teaching and Learning. Cambridge, Mass: Harvard University Press.

Bowker, G. (2006). Memory Practices in the Sciences. Cambridge, MA: MIT Press.

Bowker, G., & Star, S. L. (1999). Sorting Things Out: Classification and Its Consequences. Cambridge, MA: MIT Press.

Brophy, J. (2004). Using Video in Teacher Education. London: Amsterdam.

Brown, A., & Campione, J. (1994). Guided Discovery in a Community of Learners. In McGilly, K. (Ed.), *Classroom lessons: integrating cogntive theory and classroom practice* (pp. 229-72). Cambridge, MA: MIT Press.

Brown, S., & Eisenhardt, K. (1997). The Art of Continuous Change: Linking Complexity Theory and Time-Paced Evolution in Relentlessly Shifting Organizations. Administrative Science Quarterly, 42.

Brunner,C, Fasca, C., Heinze,J, Honey,M, Light, & Mandinach, E.(2005). Linking Data and Learning: The Grow Network Study. Journal of Education for Students Placed at Risk, 10(3), 241-267.

Bryk, A., & Hermanson. (1993). Educational Indicator Systems: Observations on Their Structure, Interpretation, and Use. Review of Research in Education, 19(3), 451-484.

Burch, P., & Spillane, J. (2005). How Subjects Matter in District Office Practice: Instructionally Relevant Policy in Urban School District Redesign. *Journal of Educational Change*, 51-76.
Burch, P. (2006). The New Educational Privatization: Educational Contracting and High Stakes Accountability. Teacher College Record, 108(12), 2582-2610.
Burch, P. (2007). Educational Policy and Practice and Institutional Theory: Crafting a Wider Lens. Educational Researcher.
Carlson, R. (1965). *Adoption of Educational Innovations*. Eugene, Oregon: University of Oregon Center for Advanced Study of Educational Innovations.
Carnoy, M., & Loeb, S. (2003). Does External Accountability Affect Student Outcomes? A Cross-state Analysis. *Educational Evaluation & Policy Analysis*, 24-4.
CCSSO/ISSC, Council of Chief State Supervisors Organization National Interstate School Leaders Licensure Consortium. (1996). *Standards For School Leaders*. Washington, DC: CCSSO.
Chen, E., Heritage, M., & Lee, K. (2005). Identifying and Monitoring Students' Learning Needs With Technology. Journal of Education for Students Placed at Risk, 10(3), 309-332.
Clay, M. (2000). Running Records: For Classroom Teachers. Portsmouth, NH: Heinemann.
Coburn, C. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. Educational Evaluation & Policy Analysis, 23(2), 145-170.
Coburn, C. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. Sociology of Education, 77, 211-244.
Coburn, C. (2005a). Shaping teacher sensemaking: School leaders and the enactment of reading policy. Educational Policy, 19(3), 476-509.
Coburn, C. (2005b). The role of non-system actors in the relationship between policy and practice: The case of reading instruction in California. Educational Evaluation & Policy Analysis, 27(1), 23-52.
Coburn, C., & Talbert, J. (2006). Conceptions of Evidence Use in School Districts: Mapping the Terrain. American Journal of Education, 467-495.
Cohen, D. K. (1995). Where is the system in systemic reform? Educational Researcher, 24(9).
Cohen, D. K. (1989). Teaching Practice: Plus qa change... In Jackson, P.W. (Ed.), *Contributing to Educational Change: Perspectives on Research and Practice* (pp. 27-89). Berkeley, CA: McCutchan.
Cohen, M., & Bacdayan, P. (1994). Organizational Routines Are Stored As Procedural Memory: Evidence from a Laboratory Study. Organizational Science, 5(4), 554-568.
Cole, M. (1996). *Cultural Psychology: A once and future discipline*. Cambridge, MA: Belknap Press of Harvard University Press.
Coleman, J. (1966). *Equality of Educational Opportunity (Coleman) Study (EEOS)*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2000: 2nd ICPSR version. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education/National Center for Education Statistics.
CCSSO/ISSC, Council of Chief State Supervisors Organization National Interstate School Leaders Licensure Consortium. "Standards For School Leaders.". Washington, DC: CCSSO, 1996.
Collins, A., Joseph, D., & Bielaczyk, K. (2004). Design Research: Theoretical and Methodological Issues. Journal of the Learning Sciences, 13(1), 15-42.
Confrey, J., & Makar, K. (2004). Using Dynamic Statistics Software to Critique and Improve Use of Data from High-Stakes Tests. In Dede, C., Honan, J. & Peters, L. (Eds.). Cambridge, MA: Harvard University Press.

Copland, M., & Knapp, M. (2006). Connecting Leadership with Learning: A Framework for Reflection, Planning, and Action. Alexandria, VA: Association for Supervision and Curriculum Development.
Creighton, T. B. (2001). Schools and Data: The Educator's Guide for Using Data to Improve Decision Making. Thousand Oaks, CA: Corwin Press.
Daft, R., & Lengel, R. (1986). Organizational Information Requirements, Media Richness and Structural Design. Management Science, 32(5), 554-571.
Davis, E., & Krajcik, J. (2005). Designing Educative Curriculum Materials to Promote Teacher Learning. Educational Researcher, 34(3), 3-14.
Diamond, J., & Spillane, J. (2004). High-Stakes Accountability in Urban Elementary Schools: Challenging or Reproducing Inequality?. Teachers College Record, 106(6), 1145-1176.
DiSessa, A. (2005). Metarepresentation: Native Competence and Targets for Instruction. Cognition and Instruction, 22(3), 293-331.
Dunbar, Kd. (1993). How Scientists Really Reason: Scientific Reasoning in Real World Laboratories. In Sternberg, R.J. & Davidson, J. (Eds.), Mechanisms of Insight. Cambridge, MA: MIT Press.
Duschl, R. (2001). Assessment of Inquiry. In Atkin, K. & Coffey, J. (Eds.), Everyday Assessment in Science Classrooms (pp. 41-59). Washington, DC: National Science Teaching Association.
Edelson, D., Gordin, D., & Pea, R. (1999). Addressing the Challenges of Inquiry-Based Learning Through Technology and Curriculum Design. Journal of the Learning Sciences, 8(3 & 4), 391-450.
Eisenhardt, K. (1989). Building Theories from Case Study Research. Academy of Management Review, 14(4), 532-550.
Eisenstein, E. (1979). The Printing Press as an Agent of Change: Cambridge University Press.
Elmore, R., Peterson, P., & McCarthey, S. (1996). The Puzzle of Organization and Practice. In *Restructuring in the Classroom: Teaching, Learning, and School Organization* (pp. 1-14). San Francisco, CA: Jossey-Bass.
Emerson, R., Fretz, R., & Shaw, L. (1995). Writing Ethnographic Fieldnotes. Chicago: University of Chicago Press.
Engeström, Y. (1987). Learning by Expanding : An Activity-Theoretical Approach to Developmental Research. Finland: Orienta-Konsultit Oy.
Erickson, F. (1996). Going for the zone: the social and cognitive ecology of teacher--student interaction in classroom conversation. In Hicks, D. (Ed.), Discourse, Learning, and Schooling (pp. 29-62). New York, NY: Cambridge University Press.
Erickson, F. (2007). Proximal formative assessment of student learning: An approach to the conduct of instruction at a university laboratory school. In Moss, P. (Ed.), National Society for the Study of Education Yearbook, Part 1: "Evidence and Decision Making". (Vol. 1,). Chicago, Ill.: NSSE : Distributed by the University of Chicago Press.
Erickson, F., & Schulz, J. (1981). When is a Context? Some Issues and Methods in the Analysis of Social Competence. In Green, J. & Wallat, C. (Eds.), *Ethnography and language in educational settings* (pp. 147-60). Norwood, NJ: Ablex.
Falk, J., & Drayton, B. (2004). State testing and inquiry based science: Are they complementary or competing reforms? Journal of Educational Change, 5(4), 344-387.
Feldman, M., & March, J. G. (1981). Information in Organizations as Signal and Symbol . Administrative Science Quarterly, 26(2), 171-186.
Fink, E., & Resnick, L. (2001). Developing Principals as Instructional Leaders. Phi Delta Kappan.
Firestone, W., & Gonzales, R. (2007). Culture and Processes Affecting Data Use In School Districts. In Moss, P. (Ed.), National Society for the Study of Education Yearbook, Part

1: "Evidence and Decision Making". (Vol. 1,). Chicago, Ill.: NSSE : Distributed by the University of Chicago Press.
Frank, K., Zhao, Y., & Borman. (2004). Social Capital and the Diffusion of Innovations within Organizations: Application to the Implementation of Computer Technology in Schools. Sociology of Education, 77, 148-171.
Frederiksen, J., & Collins, A. (1989). A Systems Approach to Educational Testing. Educational Researcher, 18(9), 27-32.
Friedman, T. L. (2005). The World Is Flat: A Brief History of the Twenty-first Century. New York: Farrar, Straus and Giroux.
Gamson, D. (2007). Historical Perspectives on Democratic Decision Making in Education: Paradigms, Paradoxes, and Promises. In Moss, P. (Ed.), National Society for the Study of Education Yearbook, Part 1: "Evidence and Decision Making". (Vol. 1,). Chicago, Ill.: NSSE : Distributed by the University of Chicago Press.
Goodwin, C. (1994). Professional Vision. *American Anthropologist*, *906*, 606-633.
Goodwin, C., & Duanti, A. (1992). Introduction. In Goodwin, C. & Duanti, A. (Eds.), *Rethinking Context: Language as an Interactive Phenomenon*. New York, NY: Cambridge University Press.
Guare, J. (Writer) & F. Schepisi (Director) (1993). Six Degrees of Separation. In A. Milchan (Producer): MGM.
Hacking, I. (1983). *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. New York, NY: Cambridge University Press.
Halliday, M. A. (1978). Language as a social semiotic. Baltimore, Maryland: University Park Press.
Halverson, R. (2003). Systems of Practice: How Leaders Use Artifacts to Create Professional Community in Schools. Education Policy Analysis Archives, 11(37).
Harris, C. J., McNeill, K. L., Lizotte, D. L., Marx, R. W., & Krajcik, J. (2003). Usable Assessments for Teaching Science Content and Inquiry Standards. *Peers Matter*, *1*(1).
Hickey, D., & Anderson, K. (2007). Situative approaches to assessment for resolving problems in educational testing and transforming communities of educational practice. In Moss, P. (Ed.), National Society for the Study of Education Yearbook, Part 1: "Evidence and Decision Making". (Vol. 1,). Chicago, Ill.: NSSE : Distributed by the University of Chicago Press.
Hiebert, J., Gallimore, R., & Stigler, J. (2002). A Knowledge Base for the Teaching Profession: What Would It Look Like and How Can We Get One? Educational Researcher, 31(5), 3-15.
Honig, M. I. (2006). Street-Level Bureaucracy Revisited: Frontline District Central-Office Administrators as Boundary Spanners in Education Policy Implementation. Educational Evaluation & Policy Analysis, 28(4), 357-383.
Iedema, R. (2003). Multimodality, resmiotization: extending the analysis of discourse as multi-semiotic practice. Visual Communication, 2(1), 29-57.
Ikemoto, G., & Marsh, J. (2007). Cutting Through the "Data Driven" Mantra: Different Conceptions of Data-Driven Decision-Making. In Moss, P. (Ed.), National Society for the Study of Education Yearbook, Part 1: "Evidence and Decision Making". (Vol. 1,). Chicago, Ill.: NSSE : Distributed by the University of Chicago Press.
Ingram, D., Louis, K. S., & Schroeder, R. (2004). Accountability Policies and Teacher Decision Making: Barriers to the Use of Data to Improve Practice. Teacher College Record, 106(6), 1258-1287.
Jackson, P. W. (1968). Life in Classrooms. New York: Holt Reinhart.
Kerr, K., Marsh, J, Ikemoto, G., 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, 496-520.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. Cognitive Science, 12, 1-48. 387-420.
Knapp, M. S. , Swinnerton, J., M, & Monpas-Huber, J. (2006). Data-informed Leadership in Education: The Wallace Foundation.
Knapp, M., Copeland, M., & Swinnerton, J. (2007). Data-informed Leadership: Insights from Current Research, Theory, and Practice. In Moss, P. (Ed.), National Society for the Study of Education Yearbook, Part 1: "Evidence and Decision Making". (Vol. 1,). Chicago, Ill.: NSSE : Distributed by the University of Chicago Press.
Krajcik, J., McNeill, K.L. & Reiser, B. (in press). Learning-goals-driven design model: Curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*
Krajcik, J., & Blumenfeld, P. (2006). Project-based learning. In Sawyer, K. (Ed.), Cambridge Handbook of the Learning Sciences. New York: Cambridge University Press.
Kuhn, T. (1970). *The Structure of Scientific Revolutions* (2nd ed.). Chicago: University of Chicago Press.
Lagemann, E. C. (2000). An Elusive Science: The Troubling History of Education Research. Chicago: University of Chicago Press.
Lampert, M. (2001). *Teaching Problems and the Problems of Teaching*. New Haven: Yale University Press.
Lather, P. (1999). To be of use: The work of reviewing. Review of Educational Research, 69(1), 2-7.
Latour, B., & Woolgar, S. (1979). Laboratory Life: The Construction of Scientific Facts. Princeton, NJ: Princeton University Press.
Latour, B. (1987). *Science in action : how to follow scientists and engineers through society*. Cambridge, MA: Harvard University Press.
Latour, B. (1990). Drawing things together. In Lynch, M. & Woolgar, S. (Eds.), *Representation in Scientific Practice* (pp. 19-68). Cambridge Mass: MIT Press.
Latour, B. (1996). *Aramis, or the Love of Technology*. Cambridge, MA: Harvard University Press.
Latour, B. (2005). Reassembling the Social: An Introduction to Actor-Network-Theory. New York: Oxford.
Lave, J., & Wenger, E. (1991). Situated Learning: Legitimate Peripheral Participation. Cambridge, UK: Cambridge University Press.
Law, J. (1992). Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity. Systemic Practice and Action Research, 5(4), 379-393.
Lee, C. (1995). A Culturally Based Cognitive Apprenticeship: Teaching African American High School Students Skills in Literary Interpretation,. Reading Research Quarterly, 30(4), 608-630.
Lee, V., & Bryk, A. (1989). A multilevel model of the social distribution of highschool achievement. Sociology of Education, 62(2), 172-192.
Lee, V., & Loeb, S. (2000). School Size in Chicago Elementary Schools: Effects on Teachers' Attitudes and Students' Achievement . *American Educational Research Journal, 37*(1), 3-31.
Leithwood, K., & Riehl, C. (2003). What Do We Already Know About Successful School Leadership?: AERA Division A Task Force on Developing Research in Educational
Lemke, J. (1998). Multimedia Literacy Demands of the Scientific Curriculum. Linguistics and Education, 10(3), 247-272.
Lemke, J. (2000). Across the Scales of Time: Artifacts, Activities, and Meanings in Ecosocial Systems. Mind, Culture, Activity, 7(4), 273-290.
Lemke, J., & Sabelli, N. (2008). Complex Systems and Educational Change: Towards a New Research Agenda. *Educational Philosophy and Theory*, *40*(1), 118-129.

Lencioni, P. (2006). Silos, Politics and Turf Wars: A Leadership Fable About Destroying the Barriers That Turn Colleagues Into Competitors: Jossey-Bass.
Lerner, L. (1998). *State Science Standards: An Appraisal of Science Standards in 36 States*. Washington, DC: The Thomas B. Fordham Foundation.
Lessig, L. (1999). *CODE and Other Laws of Cyberspace* (First ed.). New York: Basic Books.
Lipset, Troow, & Coleman. (1956). Union Democracy: The Internal Politics of the International Typographical Union. Glencoe, Il: Free Press.
Little, J. W. (2002). Locating learning in teachers' communities of practice: opening up problems of analysis in records of everyday work Locating learning in teachers' communities of practice: opening up problems of analysis in records of everyday work . *Teaching and Teacher Education*, *18*(8), 917-946.
Little, J. W. (2007). The Significance of Teachers' Classroom Accounts as Resources for Learning from and for Practice. In Moss, P. (Ed.), *National Society for the Study of Education Yearbook, Part 1: "Evidence and Decision Making"*. (Vol. 1,). Chicago, Ill.: NSSE : Distributed by the University of Chicago Press.
Louis, K. S., & Dentler, R. A. (1988). Knowledge Use and School Improvement. Curriculum Inquiry, 18(1), 33-62.
Love, N. (2002). Using Data, Getting Results: A Practical Guide for School Improvement in Mathematics and Science. Norwood, MA: Christopher Gordon.
MAISA. (2005). *2005 Data Warehousing ISD Survey http://www.michiganedusource.org/Technology/DataWarehousingSummary.pdf*.
MAISA. (2007). *2007 Data Warehousing ISD Survey http://www.michiganedusource.org/Technology/DataWarehousingSummary07.pdf*.
Mahajan, V., Muller, & Bass, F. F. M. (1990). New Product Diffusion Models in Marketing: A Review and Directions for Research. *Journal of Marketing, 54*(1), 1-26.
Mandinach, E., Honey, M., & Light, D. (2006). A Theoretical Framework for Data-Driven Decision Making. Paper presented at the American Educational Research Association, San Francisco.
March, J. G. (1987). Ambiguity and accounting: The elusive link between information and decision making. Accounting, Organizations and Society, 12(2), 153-168.
Marsh, H. (1991). Public, Catholic Single-Sex, and Catholic Coeducational High Schools: Their Effects on Achievement, Affect, and Behaviors. *American Journal of Education*, *99*(3), 320-356.
Marsh, J., Kerr, K., Ikemoto, G., Darilek, H., Suttorp, M., & Zimmer, R., et al. (2005). The Role of Districts in Fostering Instructional Improvement: Lessons from Three Urban Districts Partnered with the Institute for Learning. Santa Monica, CA: RAND Corporation.
Martin, N., & Yin, N. (1999). Beliefs Regarding Classroom Management Style: Differences Between Urban and Rural Secondary Level Teachers. *Journal of Research into Rural Education*, *15*(2).
Massell, D. (2001). The Theory and Practice of Using Data to Build Capacity: State and Local Strategies and their Effects. In Fuhrman, S. (Ed.), From the Capital to the Classroom: Standards-Based Reform in the States, Yearbook of the National Society for the Study of Education Yearbook (Vol. 100(2),). Chicago, IL: University of Chicago Press.
Maxwell, J. (2006). Literature Reviews of, and for, Educational Research: A Commentary on Boote and Beile's "Scholars Before Researchers". Educational Researcher, 35(9), 28-31.
McLaughlin, M. (2008). *Practicing community: Changing understandings (Distinguished Contribution to Educational Research Lecture)*. Paper presented at the American Educational Research Association Annual Meeting, New York, City.
Meiles, T., & Foley, E. (2005). From Data to Decisions: Lessons from School Districts Using Data Warehousing. Boston, MA: Annenberg Institute for School Reform at Brown University.

Merriam, S. (1998). -Webster. San Francisco, CA: Jossey-Bass.
Messick, S. (1989). Validity. In Linn, R. (Ed.), Educational Measurement (3rd ed., pp. 13-03). New York: American Council on Education.
Meyer, J., & Rowan, B. (1977). Institutional organizations: Formal structure as myth and ceremony. American Journal of Sociology(83), 340.
Moll, L. (1998). Through the Mediation of Others: Vygotskian Research on Teaching. In Richardson, V. (Ed.), Handbook of research on teaching (pp. 111-29).
Moss, P. (2004). The Risks of Coherence. In Wilson, M. (Ed.), *Towards coherence between classroom assessment and accountability*. Chicago, IL: National Society for the Study of Education.
Moss, P. (2005). Toward "Epistemic Reflexivity" in Educational Research: A Response to Scientific Research in Education. *Teachers College Record*, *107*(1), 19-29.
Moss, P., Girard, B., & Haniford, L. (2006). Validity in Educational Assessment. In Green, J. & Allan, L. (Eds.), *Review of Research in Education* (Vol. 30, pp. 109-62): Sage.
Moss, P., & Piety, P. (2007). Introduction: Evidence and Decision Making. In. Chicago, IL: National Society for the Study of Education.
Moss, P., & Shutz, A. (2001). Educational Standards, Assessment, and the Search for Consensus. *American Educational Research Journal, 38*(1), 37-70.
National Assessment Governing Board. (2006). Science Framework for the 2009 National Assessment of Educational Progress (Prepublication Edition). Washington, DC: National Assessment Governing Board.
National Center on Education and the Economy. (1990). America's Choice: high skills or low wages. Washington, DC: National Center on Education and the Economy.
NCLB No Child Left Behind Act. "Public Law No. 107–110, 115 Statute 1425.", 2002.
Nersessian, N. (2007). The Cognitive-Cultural Systems of the Research Laboratory. Organizational Studies, 27(1), 125-145.
Novak, J., & Gowin, D. (1984). Learning how to learn. New York, NY: Cambridge University Press.
NRC (2001). Classroom assessment and the National Science Education Standards. Washington, DC: National Academy Press.
NRC, National Research Council. (2002). Scientific Research in Education. Washington, DC: National Academy Press.
NRC, (1996). *National Science Education Standards*. Washington, DC: National Academy Press.
Ochs, E. (1979). Transcription as Theory. In Ochs, E. & Schiefflin, B. (Eds.), Developmental Pragmatics (pp. 43-72). New York City: Academic Press, Inc.
Orton, D., & Weick, K. (1990). Loose Coupled Systems: A Reconceptualization. Academy of Management Review, 15(2), 203-223.
Patton, M. (2001). Qualitative Evaluation and Research Methods (Third ed.). New York: Sage.
Pea, R., Lindgren, & Rosen. (2006). Computer-Supported Collaborative Video Analysis. Paper presented at the International Conference of the Learning Sciences, Bloomington, IN.
Pellegrino, J., Chudowsky, N., & Glaser, R. (2001). Knowing what Students Know : The Science and Design of Educational Assessment. Washington, DC: National Academy Press; National Research Council.
Penuel, W., Sussex, W., & Hoadley. (2006). Investigating the Potential of Using Social Network Analysis in Educational Evaluation. American Journal of Evaluation, 27(4), 437-451.
Petrides, & Nodine. (2003). *Knowledge Management in Education: Defining the Landscape.* Half Moon Bay, CA: Institute for the Study of Knowledge Management in Education.
Phillips, D. (2007). Adding Complexity: Philosophical Perspectives on the Relationship Between Evidence and Policy. In Moss, P. (Ed.), National Society for the Study of Education Yearbook, Part 1: "Evidence and Decision Making". (Vol. 1,). Chicago, Ill.: NSSE : Distributed by the University of Chicago Press.

Piety, P., & Palincsar, A. (2006). *"How Do We See?": Information Architecture as Theory*. Paper presented at the International Conference of the Learning Sciences, Bloomington, Indiana.
Piety, P. (2007). Learning Progressions: Systemic Considerations for Implementation. Paper presented at the American Educational Research Association, Chicago, IL.
Popham, W. (1987). The merits of measurement-driven instruction. Phi Delta Kappan.
Popham, W. J., Keller, T., Moulding, Pellegrino, J., & Sandifer, P. (2005). Instructionally Supportive Accountability Tests in Science: A Viable Assessment Option? Measurement: Interdisciplinary Research and Perspectives, 3(3), 121-179.
Provan, K, & Milward, H (1995). A Preliminary Theory of Interorganizational Network Effectiveness: A Comparative Study of Four Community Mental Health Systems. Administrative Science Quarterly, 40(1), 1-33.
Raudenush, S., & Bryk, A. (2002). Hierarchical Linear Models: Application and Data Analysis. Newberry Park, CA: Sage.
Raudenbush, S., & Willms, J. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, *20*(4), 307-337.
Rogers, E. M. (1995). Diffusion of Innovations (4 ed.). New York City: The Free Press.
Rowan, B. (1990). Commitment and Control: Alternative Strategies for the Organizational Design of Schools. Review of Research in Education.
Schon, D. (1987). Educating The Reflective Practitioner. New York: Jossey-Bass.
Scollon, R. (2001). Mediated Discourse: The Nexus of Practice. new York: Routledge.
Scott, J. (2000). Social Network Analysis: A Handbook. New York: Sage.
Silverstein, M. (1996). The Secret Life of Texts. In Silverstein, M. & Urban, G. (Eds.), *Natural Histories of Discourse*. Chicago: The University of Chicago Press.
Simon, H. (1996). *The sciences of the artificial*. Cambridge, MA: MIT Press.
Smith, C., Wiser, M., Anderson, C., & Krajcik, J. (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and the atomic molecular theory. Measurement: Interdisciplinary Research and Perspectives, in press.
Spillane, J. (1998). State Policy and the Non-Monolithic Nature of the Local School District: Organizational and Professional Considerations. American Educational Research Journal, 35(1), 33-63.
Spillane, J. (2004). *Standards Deviations: How Schools Misunderstand Education Policy*. Washington DC: Harvard University Press.
Spillane, J. (2006). Distributed Leadership. Hoboken, NJ: Jossey-Bass.
Spillane, J., & Burch, P. (2004). Policy, Administration, and Instructional Practice: "Loose Coupling" Revisited (No. WP 03-04). Evanston, IL: Northwestern University Institute for Policy Research.
Spillane, J., Camburn, E., Lewis, G., & Stitziel-Pareja, A. (2006). *Taking a Distributed Perspective in Studying School Leadership and Management: Epistemological and Methodological Trade-offs*. Paper presented at the American Educational Research Association, San Francisco, CA.
Spillane, J., Diamond, J., & Jita. (2003). Leading instruction: The distribution of leadership for instruction. Journal of Curriculum Studies, 36(1), 3-34.
Spillane, J., Halvorson, R., & Diamond, J. (2003). Towards a theory of leadership practice: a distributed perspective. Journal of Curriculum Studies, 36(1), 3-34.
Stake, R. (1978). The Case Study Method in Social Inquiry. Educational Researcher, 7(2), 5-8.
Star, S. L., & Griesemer, J. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology,. Social Studies in Science, 19(3),

Streifer, P. A. (2002). Using Data to Make Better Educational DecisionsUsing Data to Make Better Educational Decisions. Lanham, MD: Scarecrow Press.
Successline, I. (2002). *Designing and Using High-Quality Paper-and-Pencil Tests*. Suffolk, VA: Successline Publications.
Supovitz, J., & Klein, V. (2003). Mapping a Course for Improved Student Learning: How Innovative Schools Systematically Use Student Performance Data to Guide Improvement. Philadelphia, PA: Consortium for Policy Research in Education, University of Pennsylvania.
Swales, J. (1998). Other floors, other voices: a textography of a small university building. Matwah, NJ: Lawrence Erlbaum Associates.
Swales, J. (2004). *Research Genres: Explorations and Applications*: Cambridge University Press.
Thorn, C. (2001). Knowledge Management for Educational Information Systems: What Is the State of the Field? 9, 47.
Tufte, E. (2001). The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press.
Tushman, M., & Scanlan, T. (1981). Characteristics and External Orientations of Boundary Spanning Individuals. The Academy of Management Journal, 24(1), 83-98.
Tversky, B. (1981). Distortion in Memory for Maps. *Cognitive Psychology, 13*, 407-433.
Vygotsky, L. (1985). Thought and Language. Boston, MA: MIT Press.
Wade, S., & Zone, J. (2000). Creating Inclusive Classrooms: An Overview. In Wade, S. (Ed.), *Inclusive Education: A casebook and readings for prospective and practicing teachers* (pp. 3-28). Matwah, NJ: Lawrence Erlbaum Associates.
Wayman, J., & Stringfield, S. (2006). Technology-Supported Involvement of Entire Faculties in Examination of Student Data for Instructional Improvement. American Journal of Education, 549-571.
Weatherley, R., & Lipsky, M. (1977). Street-Level Bureaucrats and Institutional Innovation: Implementing Special-Education Reform. *Harvard Educational Review*, *47*(2), 171-197.
Weick, K. (1976). Educational organizations as loosely coupled systems. Administrative Science Quarterly, 21(1).
Weick, K. (1995). Sensemaking in organizations. Thousand Oaks, CA: Sage.
Weisenbach, T. (2007, October 12 2007). Consolidation of school services focus of new state law. *The Huron Daily Tribune*.
Weaver-Hightower. (2008). An Ecology Metaphor for Educational Policy Analysis: A Call to Complexity. *Educational Researcher*, *36*, 153-167.
Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, MA: Cambridge University Press.
Wenger, E. (2000). Communities of Practice and Social Learning Systems. Organization, 7, 225
Wertsch, J. (1998). Mind as Action. Oxford, UK: Oxford University Press.
Wilson, M., & Bertenthal, M. (2005). Systems for State Science Assessment. Washington, DC: National Academy Press.
Winner, L. (1986). Do Artifacts Have Politics?. In *The Whale and the Reactor* (pp. 19-39). Chicago: University of Chicago Press.
Yin, R. K. (2002). Case Study Research. Design and Methods (Third ed.). California: Sage.
Yinger, R. (1990). The Conversation of Practice. In Clift, R., Houston, R. & Pugach (Eds.), Encouraging reflective practice: An examination of issues and exemplars. new York: Teachers College Press.
Young, V. (2006). Teachers' Use of Data: Loose Coupling, Agenda Setting, and Team Norms. American Journal of Education, 521-548.