Educational turbulence: The influence of macro and micro policy on science education reform

Abstract

Enactment of federal educational policy has direct implications for states and local school districts across the nation, particularly in the areas of accountability and funding. This study utilized constructivist grounded theory to examine the impact of policy on science education reform in a large, urban school district over a five-year period. The existence and interaction between macro and micro, explicit and implicit policies created educational turbulence. Findings further extend upon Fullan’s (2006) change theory adding high-stakes accountability as a prevalent distractor issue and the need for quad-level, rather than tri-level engagement in reform. Suggestions for addressing educational turbulence are provided.
As we anxiously seek to remedy the broken U.S. educational system, our children and teachers live in the aftermath of educational turbulence. In an attempt to generate reform, federal legislation has mandated higher standards and increased accountability for states to implement in reading and mathematics. Unfortunately, this resulted in fifty versions of academic standards and low-quality assessments created on a shoestring budget. As a result, state-level administration of woefully under resourced policy has created a climate of disarray that must be addressed for science reform to be realized. Educational turbulence is “interplay of external variables that directly influence school reform” (p. xv) including, primarily, the enactment of macro and micro educational policy (Author, 2011a).

Enactment of federal educational policy has direct implications for states and local school districts across the nation, particularly in the areas of accountability and funding. Over the past decade, federal agencies and government officials have called for educational improvements focused on preparing students with skills required to be successful in the 21st century (Committee on Prospering in the Global Economy of the 21st Century, 2007). Concurrently, the No Child Left Behind Act (NCLB) of 2001 legislation sparked renewed high expectations for all students to learn, with a goal to “close the achievement gap between nonmainstream and non-minority students and between disadvantaged children and their more advantaged peers” (No Child Left Behind Act of 2001, 20 U.S.C. § 1001).

NCLB included mandates for state-developed accountability measures. By 2014, all students are expected to be proficient in reading and mathematics. Districts and schools that fail to make progress toward annual proficiency goals run the risk of having reduced federal funding
and potential state administrative takeover. Many have argued that accountability policy has
done more harm, particularly for diverse students, than good (Skerrett & Hargreaves, 2008).
Secretary of Education Arne Duncan also expressed looming doubts regarding federal NCLB
policy. In fact, Duncan announced in August 2011 plans to issue waivers to release states from
NCLB reprimand that agree to enact President Obama’s preferred education reforms, including
sanctions on teachers (Neill, 2011). The failure of NCLB to produce desired outcomes is
primarily due to lack of provisions for capacity building or necessary infrastructure and state
control of developing academic standards and assessments of progress (Fullan, 2010). Moreover,
NCLB implementation has produced unintended consequences for K-12 classrooms – creating
educational turbulence and associated diminished quality of teaching and learning of science and
other disciplines in schools.

As alluded to previously, state development and control of academic standards and
assessment systems have resulted in a plethora of variations of benchmarks and testing
instruments across the U.S. Undoubtedly, financial concern regarding the arduous task of scoring
assessments for millions of students each year was warranted. As a result, existing state tests in
science and other content areas are multiple choice, focused on recall of isolated facts and absent
of any real-world application or problem solving (Author, 2009; Goldston, 2005). State standards
for each discipline are broad—covering much more content than in many other countries
(Darling-Hammond, 2010), though state assessments measure performance on only a small
subset consisting of around 25 items for the 3-year grade band span.

Educational turbulence is most prevalent at the local school district level due to the fact
that macro-level (federal and state) policies result in creation of micro-level (district and school)
policies designed to produce desired outcomes. Often, policies at the local level are created in a
kneejerk fashion, failing to consider impact on students or support needed to realize necessary improvements in teacher quality. Fullan (2010) argued that current strategies for reform in the U.S. have caused “dysfunctional outcomes” (p.24). Furthermore, as Hess (1999) argued, “reform efforts are more heavily influenced by political pressures than by educational considerations” (p. 5).

Darling-Hammond (2010) characterized the detrimental impact of macro-level accountability on micro-level school policy, as resulting in “substantial instructional time on exercises that look just like the test items, reverting to worksheets filled with multiple-choice questions and drill based on recall and recitation” (p. 71), which diminishes high-quality teaching. Another implication of the increased policy focus on assessment is the reduction and often elimination of instructional time for subjects that are not tested including science (e.g., Author et al., 2010; Lee & Luykx, 2005). Moreover, the powerful influence of high-stakes accountability and existence of educational turbulence within local educational systems has translated into unintended detrimental consequences for many of our nation’s children.

In the midst of mandates, the call for improving educational systems, including science teacher quality, have been prominent. In *The Flat World and Education: How America’s Commitment to Equity will Determine Our Future*, Darling-Hammond (2010) suggested that teacher quality is the most important influence on student learning. Interestingly, federal accountability mandates focused on improving student performance have failed to include funding to improve instruction, though many have argued that teacher change in practice is the most challenging part of reform (e.g., Anderson, 2002; Desimone, 2002; Lee & Luykx, 2005). Policymakers have implied that educators and schools could be reformed to deliver new ways of learning including the integration of 21st Century Skills without support. Transforming education
is much like transforming practice in other professions (i.e., medical field, technology); training and support are needed to improve teacher quality (Author, 2011b; Desimone, 2009; Lee & Luykx, 2005; Putnam & Borko, 1997).

Prior to NCLB, most reforms targeted modifications of systems and policies without regard to changing actual classroom practice, thus producing unsuccessful outcomes (Cohen, 1995; Elmore, 1996; Hess, 1999). Interestingly, research on transforming teacher quality has grown particularly in identifying necessary conditions for change to take place. Desimone (2009) conducted a review of literature on effective professional development and identified characteristics that are key to producing systemic and sustained change in teacher practice. Desimone organized the components into a core conceptual framework that she considered “critical to increasing teacher knowledge and skills and improving their practice, which hold promise for increasing student achievement” (p. 183). Components of the framework include content knowledge focus, actively learning opportunities with new content and skills, coherence of new strategies with school/district initiatives, duration of support sustained over time, and collective participation of many or all teachers from same school (Desimone, 2009). As the knowledge base on educational reform and improving teacher quality has grown over the past decade (e.g., Author et al, 2007a, 2007b; Author & other, 2010; Loucks-Horsley, Love, Stiles, Mundry, & Hewson, 2007; Putnam & Borko, 1997), it has become more evident that traditional professional development formats do not result in sustained change in practice. Professional development linked to school and district reform initiatives have the best opportunity to transform educational practice systemically. However, since enactment of NCLB few attempts have been made to explore the ability of effective teacher quality programs to achieve systemic reform (Author et al, 2007; Desimone, 2009).
Furthermore, there continues to be a dearth in the literature on the ability of effective programs to achieve intended results in a climate of high-stakes accountability (Desimone, 2009; Lee & Luykx, 2005; Southerland, 2011). Currently, no states within the U.S. have achieved whole-system reform through implementation of any initiatives (Fullan, 2010). It is more critical now than ever to examine successful teacher quality improvement programs and determine how collective capacity is built within climates of educational turbulence to achieve educational reform.

The National Center on Education and the Economy (2008) reported,

The core problem is that our education and training systems were built for another era, an era in which most workers needed only a rudimentary education. It is not possible to get where we have to go by patching that system. We can get where we must go only by changing the system itself. (p. xxv)

Change is often thought of as a long, tiring process; however, Fullan (2010) stated that reform will produce results in a reasonable time if the appropriate strategies are implemented. Hess (1999) argued that reformers should stop chasing the elusive “silver bullet” and focus on “understanding why urban school systems engage in reform and why nearly every reform produces disappointing results” (p. 6). Given the significant investment of funding by federal government and other agencies to transform practice, further research and theoretical explanations are warranted to better understand and address the influence of policy on reform.

Theoretical Framework – Change Theory

The complexities of educational reform have been discussed throughout the literature as hundreds of change-based programs are funded and implemented each year in schools across the U.S. Reformers continue to grapple with transforming the system—often one school or district at
EDUCATIONAL TURBULENCE

a time—seeking better ways of understanding the influences that facilitate or restrain educational change. Michael Fullan’s work in this area has been extensive, including a change theory (Fullan, 2006) with “seven core premises that underpin our use of change knowledge” (p. 8). Grounded theory approaches typically do not start with a theoretical frame from the beginning of the study (Glaser & Straus, 1967). Similarly, the analysis of data in our study warranted the inclusion of change theory as a compelling theoretical lens. Additionally, our grounded theory of educational turbulence extends and refines current understandings of change theory.

The first premise of change theory is motivation, which includes individual and collective motivation to change practice. Fullan said that motivation and engagement in reform are key to producing change. The second premise of change theory is capacity building, described as “any strategy that increases the collective effectiveness of a group to raise the bar and close the gap of student learning” and “the more one invests in capacity building, the more one has the right to expect greater performance” (p. 9). Learning in context was Fullan’s third premise, which focuses on acquiring new knowledge and strategies within the local school context. The fourth premise addresses changing “the larger context” to ensure and sustain change, which is essentially infrastructure building within districts and motivation to stay the course. Within the larger context it is essential to be proactive in addressing distractors, or variables within the system that favor the status quo. Distractor issues include union and collective bargaining issues, unnecessary bureaucracy, and managerial issues.

The fifth premise is having “a bias for reflective action,” (p. 10) including a shared vision and ownership of reform. Fullan explained that behavior and changes in practice often happen before beliefs are transformed. Reflection is essential to facilitate the change process. The sixth premise of change theory is tri-level engagement, an essential component to reform that includes
the school and community, the school district, and the state as key components. Tri-level engagement is “pursuing strategies that promote mutual interaction and influence within and across the three levels” but not necessarily alignment. Fullan said that alignment is “a static unachievable goal” (p. 11). Persistence and flexibility comprise the seventh premise. Fullan noted that the previous “six premises are complex to manage and must be cultivated over time, including bumpy cycles, a strong resolve is necessary to stay the course” (p. 11).

**Methods**

This study utilized grounded theory to examine the impact of policy on educational reform in an effort to generate substantive theory (Glaser & Strauss, 1967) to better understand turbulence. This study emerged out of a larger case study of fidelity of implementation of two reform-based improving science teacher quality professional development programs funded by the Institute of Education Sciences within a large, low-performing, urban school district in the southwest United States.

The study spanned 5 years (3 years for the middle school focused program and 2 years of the elementary implementation) from 2005-2010 and involved 60 science teachers (10 middle school and 50 elementary teachers) and 9 administrators (8 building principals and 1 district administrator). Within less than a month into implementation of the first year (2005) external variables were prevalent and influenced participants and the program. The size of the federal government’s investment was considerable, as was the need to uncover potential influences that facilitated or impeded the reform efforts. Therefore, a grounded theory study was implemented with the purpose of exploring the influence of policy on large-scale implementation of science education reform in this district.
Strauss and Corbin’s (1998) approach to grounded theory research was a primary guide for this study. Constant comparison of data was conducted as described by Harry, Sturges, & Klingner (2005), as it was collected over the 5-year period, “in which the researcher moved back and forth among the data and gradually advanced from coding to conceptual categories, and thence to theory development” (p. 5). Specifically, this included analysis of each transcript as collected (including member checking) and coding of data constantly as it was generated to develop initial codes and identify emerging categories and themes. The first step to data analysis was open coding of all data collected with codes emerging from the data. As each incident or action was coded, it was compared with previous events to begin to “generate theoretical properties of the category” (Strauss & Corbin, 1998, p. 106) through determining if new data referred to previous incidents or actions and should be coded similarly. Axial coding was the second step (Strauss & Corbin, 1998), where coded data were clustered and categories were formed. The third step in coding of data was selective coding where themes were constructed based on how categories relate to each other and were grouped accordingly. According to Henry, et al. (2005), “It is in seeking the interrelationships between these themes that the researcher begins to build a theory” (p. 5). Most importantly, evidence that both supports and conflicts the themes are considered to further refine the theory, and integration of both is essential to consistency of the theory (Strauss & Corbin, 1998).

To organize the long duration and large amount of data for this project, the study was broken into phases. Across the entire project, over 780 contact hours were invested in working with participants within classrooms and in professional development sessions by the lead researcher and program staff, increasing the rapport with participants and, arguably, increasing the validity of data collected. Administrators participated in formal and informal conversations
regarding the program each year, including 3 total hours of time spent with each principal and over 20 hours with the district administrator.

Phase 1

Phase 1 took place across 2005-08 during the middle school program implementation and consisted of 540 hours working with teachers in the classroom and professional development setting. This phase included six interviews with 10 participating teachers and 24 focus group sessions. All interviews and focus group sessions were digitally recorded and transcribed. Open coding was conducted on all data as it was collected relating to external influences on implementing reform by the lead researcher. Initial codes derived from the first 2 years’ worth of extensive interviews, and focus group sessions were clustered into similar groups that became categories of external policy influence for middle school teachers. Some examples of initial codes emerging from the data were number of class periods taught, number of subjects taught, time for individual planning, rigidity of pacing guide, and quality of district assessments.

Reliability for initial coding was ensured through a process where the lead researcher and research staff coded a sample of interviews independently (one for each participant) and then met to share and discuss coding schemes. This process was repeated with a focus group sample resulting in consensus on coding. Axial coding was also included in Phase 1, where open codes were grouped into categories by the lead researcher. Axial codes were discussed by the research team extensively, which resulted in more clarity and refinement of initial codes and axial codes.

In the third and final year of the middle school project, deeper probing into the established categories took place through open-ended questioning relating to the impact of policy on their teaching in the final focus group session. Responses from participants collectively confirmed the established categories. Administrators conducted one formal interview about their policies and
decision-making that was coded by the lead researcher and verified by one project staff member. Categories for administrator data were determined by the lead researcher and project staff collectively and corresponded with similar categories emerging out of teacher participant data.

**Phase 2**

Phase 2 occurred during the first year of the elementary school project (2008-09) and consisted of 120 hours working in classrooms and with teachers in professional development sessions. When the 50 elementary teachers came on board in this phase of the project, it was important to revisit initial coding and categories to see if any codes were not present in the data for the elementary teachers or if new codes emerged from this group that were not prevalent for middle school teachers. Two interviews per participant (100 total) and three focus group sessions were coded in the same manner as with the first phase of the project for the middle school teachers. Coding of the data revealed elementary teachers’ experiences were similar to those reported in the middle school data. Only one additional category was added to the list. Again, administrators were interviewed and informal conversations took place across the program, which were coded by the lead researcher and represented repetition of codes and categories from Phase 1. At the end of Phase 2, all codes and categories had reached saturation and repetition.
Phase 3

The next step in data analysis was taken at the beginning of the second year of the elementary project (2009-10), which included 120 contact hours. Selective coding was used to develop themes that would encompass the categories in Phase 3. The themes for our study were the various areas of policy that influenced implementation and success of the reform project from both the teacher and administrator perspectives. The lead researcher and program research staff met and discussed the various categories and how the issues and topics could be made concise and to the point in a theme for each. Seven themes emerged from the group as the primary areas portraying the arguments revealed in the data. The selective codes included (a) mandated accountability, (b) funding limitations, (c) mandated curriculum and instruction, (d) personnel, (e) scheduling issues, (f) learning environment quality, and (g) lack of community engagement.

Once the themes were developed, they were tested in four ways. First, one of the teacher focus group sessions was devoted to discussion of the influence of each of these areas of policy on teacher practice. We also added a question to our end-of-year teacher interview that asked specifically, “Are there other policies or practices that the school or district have that impact your teaching?” This question served as a followup to a standard question on the influence of assessment on their teaching. In response to this question, participants were able to articulate across the group issues that fit within the established themes and associated categories. Informal open-ended conversations were conducted with administrators regarding the influence of educational policy mandates on their district and school policy. The themes of these conversations were also repetitive of earlier coding and categories and aligned with established themes.
Last, the themes were applied to the other two teacher focus group sessions and three classroom observations conducted that year, revealing prevalence of themes in discussions not directly focused on eliciting them and influences of many themes on the classroom setting. In addition to testing of selective codes, data were triangulated using teacher and administrator interviews, teacher focus groups, classroom observations of teachers and artifacts, such as district communications and professional development meeting notes.

Phase 4

Phase 4 took place at the conclusion of implementation of the project intervention (2010-11) and consisted of generation of substantive theory based upon the detailed findings in our 5-year study. Data were collected and analyzed concurrently across the project, allowing for consistency checks within and between data sources for individuals and exploration of any ambiguities emerging from the data. In this phase, established themes emerging from our selective codes are identified as central and component concepts within our grounded theory.

Sources of Educational Turbulence

In our exploration of the sources of educational turbulence for teachers within a large urban district engaged in systemic educational reform, seven themes (policy areas) emerged from the data: accountability, funding, curriculum and instruction, personnel, scheduling, learning environment, and community engagement. Due to the emphasis of this reform project being within a large school district, data relating to micro-level policy were much more prevalent than macro-level policy. However, macro-level policy was discussed by participants as well as the associated influence on micro-level policy actions. In our discussion of data, the central concept of accountability is presented, followed by component concepts and our grounded theory.
diagram of educational turbulence. Pseudonyms will be used for the reporting of all names of schools and teachers in this manuscript.

**The Reach of Accountability Policy**

The formidable challenge and emerging central concept in this study was the impact of district- and state-mandated assessments on participant ability to successfully implement reform-based practice. In fact, district administrators argued that macro-level accountability policy had become the driving force behind the goals and priorities of the district and had permeated every facet of policy. As Mr. Muncy, district administrator, said, “Raising our scores on state testing in reading and mathematics are our primary goals for survival, and we have aligned our resources and devoted more instructional time to try and make this happen.” In this state, high-stakes assessments were administered each year for reading, mathematics, and science beginning in third grade (rather than testing at the end of grade bands) going through eighth grade. High school assessments were also in place corresponding to coursework exit exams. Clearly, reading and mathematics performance were more critical to the district in elementary and middle grades, as Mr. Muncy said, similar to many other district priorities across the nation responding to No Child Left Behind accountability measures that focus only on those two subjects.

The district addressed the goal of improved student performance on state assessments through new micro-level policies, including mandated district pacing guides and teacher developed end-of-unit assessments for elementary and middle grades. The rationale behind these policies was addressing high student mobility across the district and having a diagnostic tool in place to enable teachers to provide further instruction for students to attain mastery of concepts.
Arguably, this plan was flawed from the beginning, because the pacing guide and diagnostic assessments were inserted into a system that was not set up to succeed.

One of many issues was the large amount of curriculum packed into each grade level in every subject area. Most teachers failed to cover the required content of the curriculum prior to the administration date for assessments set by the district. Teachers were forced to administer the assessment regardless of whether they had completed the instructional sequence. Students then experienced repeated failure across the academic year (average of 15 total assessments), because they had not been taught the tested concepts. In addition, the ability of district assessments to serve as diagnostic tools was undermined, because many teachers chose not to remediate. They instead chose to teach concepts they had not covered previously. Further, pacing guides and district schedules became obsolete except for stopping to review for an assessment and attempting to cram in vocabulary the day before the test. Mrs. Haring, a sixth-grade teacher in the program shared her frustrations regarding the pacing guide and the impact it had on her:

Well, I have been reprimanded again and again for going too slow—for letting the students ask too many questions. [sigh] I have, it’s all about pacing and my pace is too slow. I’ve been asked over and over again why my pace is too slow. Because I encourage students to ask questions, and we discuss their questions (Interview, 2009).

In addition to the unrealistic schedule and mandated testing days, there was another considerable flaw to district assessment: the fact that testing instruments were developed by elementary and middle school teachers who were not content experts and by district personnel who were not teaching the content at all. Further, the items were poorly constructed multiple-choice questions, full of inaccuracies and misconceptions. For example, one item on a sixth grade assessment asked students to examine a diagram of the solar system, which was labeled
“Our Universe”. On a fourth grade test, students were to examine a diagram of ripples in the sand in the ocean floor and were asked, “What type of fossil is this?”. Mr. Roberts, a seventh-grade middle-school teacher expressed his displeasure regarding district unit assessments and the impact on his ability to teach creatively.

They tell us at the district level, the administrative level, block tests are so important. We spent all this money on this new machine….to grade these things for some kind of feedback. If you ask me, it’s all crap. The whole freaking thing is all crap. I’d like to just throw them all out and teach the way I feel deep down inside it should be done and not what they’re asking me to do to have kids memorize things they will forget the day after the test (Interview, spring 2008).

Participants in this study expressed legitimate concerns regarding the administration of assessments and how the concepts should be taught (i.e., should they teach the correct answers or the answers expected by the test authors?). They were also concerned that test scores were required to count as a grade for students. Mr. Wyan, a fourth-grade teacher, expressed his frustration regarding the lack of consistency and format of unit assessments:

I wish it would be more standardized. Meaning now for block one there might be 10 questions, and for others there were like 28. The tests were very specific I felt like and not necessarily a good indicator of what, how much knowledge my students actually gained through the course of those 6 weeks (Interview, 2010).

It is important to note that district assessments were constructed to model state assessment released items, which also focused on recall and memorization of vocabulary definitions rather than on real-world application of concepts. Mrs. Franklin discussed her disapproval of the system of testing: “I don’t think just repeating it and being a bunch of parrots
necessarily made them learn a thing.” Moreover, a system set up to mimic the state assessments and prepare students to improve performance through practicing and “pacing for all” was failing to produce anything more than teacher and student frustration and anxiety.

**Component Concept: Funding Policy**

This study was conducted in a state that had one of the lowest per-pupil funding levels in the nation: less than $6,000 per student compared to over $10,000 as a national average. The macro-funding policy in this state has implications directly relating to all areas of micro-policy that will be discussed further in those areas of our paper. Obviously, the reduced funding for education within this state translated into low-quality state assessments (multiple-choice format) that are less expensive to score than more authentic open-response instruments (which would allow for more insight into student learning and their ability to apply knowledge rather than only recall it).

Another implication of the insufficient funding for education in this state is extremely low teacher salaries (starting around $26,000) with average salary of $40,000, ranking 45th nationally. Many teachers worked second jobs to make ends meet. Second jobs coupled with little to no time for instructional planning resulted in lack of flexibility for teachers to be innovative on their own time and to be available for additional student support after school (which is unpaid by the district).

Undoubtedly, school districts are required to operate within their funding availability. Often, administrators must make difficult decisions that are not aligned with what is in the best interests of children and teachers. The district in this study chose to close a smaller, low-performing middle school and shift boundaries within the district to move students to the three remaining schools. Class size in the schools increased to over 35 students and, in many cases,
nearly 40 per class. Many teachers did not have enough seats in their crowded rooms to accommodate students.

Funding for teacher supplies, professional development, and field trips was limited and prioritized based upon district goals. Because the district was a low-performing, large urban district with a large population of Hispanic and English language learner students, many grant resources were available, including the two research grants we implemented to address their needs. Unfortunately, as many participants noted, the district was anxious for improvement and accepted almost all forms of help that brought funding to the district regardless of focus area. As a result, multiple programs were implemented concurrently. Reading First is one of the many programs that had a long tenure in the district, as well as a state-level mathematics program, School Improvement Grants, and others. Overburdened teachers were required to participate in many funded programs—most not implemented long enough to determine the impact before another program was tried. The district allocated internal funding to add an English as a Second Language endorsement (ESL) and required all teachers to complete it within a short time frame. ESL classes were scheduled through the district and taught in the evening after school.

A large investment was made in creating a science magnet elementary school, including a new building equipped with the latest technology and housed in one of the most needy areas of the district. Unfortunately, the facility was the only aspect that changed, and student performance remained low. One of the new teachers who joined our program the second year was a new hire at the science magnet school. Miss Vaughn reported, “I kind of laughed after I got here that they even tell people we are a science school when we don’t have a science time in our schedule.” She elaborated that it was up to individual teachers to find time for science and social studies if they chose to integrate it.
Component Concept: Curriculum & Instruction

Many implicit and explicit policies within the district were identified in regard to how and what to teach. The district had a large population of Hispanic students primarily from Mexico, many of which were identified as English language learners. The district had no pull-out support for English language learners within this district. Its policy was to have all teachers ESL endorsed, removing the need to have full-time ESL staff. Teachers decried the constant struggle they experienced teaching under these parameters while working with children with low literacy levels. This policy was unrealistic at best—expecting teachers to do pull out and make accommodations in overcrowded classrooms with no support and short, 45-minute class periods in middle school.

At the elementary level, mandated instructional time of 3.5 hours daily for reading and writing and 1.75 hours for mathematics (totaling 5.25 hours) left little remaining time in the day to teach other subjects or conduct interventions. Some elementary teachers said that they shuffled students to group all English language learner students in one classroom, where an assigned teacher would work with them while others did other remediation and enrichment during the day. The building administrator mandated this informal system be abandoned after only a year. According to many teachers, the policy required support of only English language learner students during the accommodation time—prohibiting inclusion of other special needs students who could benefit from support. The biggest challenge for teachers was what to do with the other students during this time, as Mr. Bryan, a fifth-grade teacher discussed: “I have times when I’m working with my ESL students and one of the kids that is lower asked if he could come over and be a part of it with us. Because of the law, I have to tell him no.”
Teachers also reported that enrichment activities for students who were on grade level were almost nonexistent. Mr. Bryan expressed his feelings in an interview relating to the increasing demands of the district for accommodations of English language learner students:

I’m getting tired of them telling me all the stuff I have to have hanging on my walls and all the different ways that I have to teach….and like I think it’s just gotten pathetic in so many ways. They come in and say you have to do this for ESL, and you have to put this on your wall….and part of me is just like, “Stop trying to tell me what to have on my walls and just let me teach.” (Interview, 2011)

An example Mr. Bryan alluded to was the district mandate that posters were to be displayed in classrooms relating to English language learner accommodations. However, the font size on the posters was too small to be read from more than one foot away, as Mr. Bennelton, a sixth-grade teacher joked, “Now we have something on our wall that no students can read, so it makes our English language learners feel better.”

The district evaluation system for elementary school teachers emphasized the use of direct instruction in the teaching of all subjects. This implicit mandate was at complete odds with the reform-based program our program was implementing within the district, as Mrs. Howe, a sixth-grade teacher shared:

I’m so excited to get away from the direct instruction and the explicit lesson plan. I don’t think that is the way to teach critical thinking skills, and that is what our students lack. They have been told so much what to think that they don’t know how to think on their own. (Focus group session, 2010)

Teachers who had transformed their beliefs about how to teach and wanted to incorporate more inquiry and the 21st Century Skills Framework (Partnership for 21st Century Learning,
were penalized during evaluations of their performance by administration. Interestingly, all building and district administrators signed on to fully support our program prior to their school being selected to participate and receive over $30,000 in materials, but district teaching evaluation rubrics did not reward the pedagogy we emphasized.

District prioritization of reading and mathematics drove the purchase of new curriculum. The implementation of new curriculum was problematic for two reasons. First, teachers were provided no professional development related to using the new materials. Second, the district changed curricular materials twice during our 5-year program—providing little time to determine impact on student learning or fidelity of implementation of curriculum by teachers. Mr. Hall described frustration regarding lack of support for teachers.

I think the reading program. I mean, well they really just hit us with it. You know, it wasn’t like training to where we could come in and feel a little comfortable, you know. We were kind of thrown into it, (Focus group session, Spring 2010)

**Component Concept: Personnel**

Expectations for teacher performance within this district were high despite minimal investments to ensure success. Teachers at many schools had teachers who were assigned two grade levels within their teaching load requiring multiple preparations within the school day. The burden for middle school teachers was not as severe as for elementary teachers, as they had a preparation period each day. Elementary teachers received no planning or preparation time during the school day, although they had preparations for all subjects taught in self-contained classrooms. Many schools due to enrollment numbers had at least one teacher who taught two grade levels combined within the same self-contained classrooms. Not only did these teachers
have to teach all subjects, but they also had to determine how to teach content of two grade levels within 1 year.

Professional development opportunities offered through the district were limited to the required ESL endorsement coursework. All newly hired teachers were required to complete this coursework along with attending new teacher orientation meetings. Existing teachers were required to complete the endorsement within a period of time to secure their jobs. Teacher had no choice in professional development options, and most options offered were funded by grants.

The teacher contract was up for renewal at the end of our study, and due to the reduced bargaining capacity of unions in that state, the superintendent drafted the contract and gave teachers ultimatums to either sign the contract or lose their jobs. The new contract contained more provisions for linking job security to performance on state assessments, thus increasing pressure to teach to the test, as Mrs. Haring reported.

They’re doing learning marks this year where they come through and they have a checklist and they can come through at any time and you need to be showing explicit instruction, following plan models, and they have different requirements at different times of the week and want to see these things on your wall (Focus group, 2010).

Component Concept: Scheduling

As alluded to in our presentation of data up to this point, scheduling policy within this district provided little room for planning, individualization, and collaboration with other professionals. At the middle school level, despite individual planning time within the school day, rigid 45-minute class periods provided no opportunities for extended learning and no periods built into the schedule or contract for meeting and collaboration with other teachers. Mrs. Fredricks, an eighth-grade middle school teacher said that her planning period “is often used to
hold special education student meetings and grade the mound of 220-plus papers I collect each
day from students.” Additionally, the short class periods and large class sizes combined for fewer
opportunities to individualize, given that middle school teachers had seven class periods a day,
totaling over 200 students. Elementary teachers received no planning time within the daily
schedule or time to meet and collaborate with others. Mrs. Wyan, a fourth-grade teacher, shared
her concerns regarding lack of time to collaborate and plan after learning the value in doing so
during the three professional development days provided by the grant program.

Well, meeting together as a team. We could do a lot better if we got more time to plan
other than grant time. But we don’t get time to plan because we are always with our class all day
long (Interview, 2009).

The struggle with trying to teach the entire curriculum within the school day coupled with
mandated time for reading and mathematics, left little time for teaching other subjects (science,
social studies, health, physical education, music, art) as Mrs. Pennick, a fourth-grade teacher,
reported:

I’m going to be totally honest with you. I think that a lot of it is lip service. I mean, it
sounds good, it feels good to say science is important. We need to teach science. But so
far, they have not said one word about what is going to go. There are just so many hours
in the day. So until somebody higher up says you have permission to do less of such and
such to do science, it will not occur. It will be scheduled and not happen (Interview,
2009).

**Component Concept: Learning environment**

Policies regarding the learning environment, specifically materials, curriculum,
technology, instructional support, and out-of-school experiences, if included, were aligned with
the focus on mathematics and reading across the district. When we began our work in the district, there was no science curriculum in place at the elementary level, which meant only isolated pockets of science were being taught. Similarly, materials across the board were few and far between. The middle school science teachers were doing experiments in canning jars due to lack of supplies and equipment for a required subject that was getting equal instructional time during the day. Students struggled to memorize the pictures of beakers and other laboratory equipment from worksheets so that they may identify them correctly on state assessments. Technology was inequitably distributed. Newer buildings (for example, the science magnet school) had a plethora of technology tools, including interactive white boards and student responders. Other schools, including middle schools, had one or two computer labs that were scheduled for keyboarding classes and difficult to gain access to. Some teachers were selected near the end of the program to participate in a grant focused on integration of technology that enabled them to get a class set of desktop computers. Most middle school classrooms had one computer and elementary schools had three to five. Support for integration of anything new across the district was problematic.

Three years into the program many of the science magnet school teachers had not used their student responders at all due to lack of training on how to integrate them into their teaching.

The district employed approximately five instructional coaches, who served 13 elementary schools, four middle schools, and two high schools. The work of the coaches was focused on reading, mathematics, and classroom management primarily. The district had no funding for field trips, even if there was no cost of the experience other than transportation. Seating for students and physical space was also a challenge for teachers in the program. At the same time the district was building new school buildings equipped with technology—other schools, including one middle and one elementary, operated in open-classroom buildings from
the 1970s that used only temporary walls without doors to hold instruction in. At the end of the program, the district added a wing to the middle school for science classrooms due to a strong building leader who advocated the need for appropriate facilities for teaching chemistry. The district moved this principal out of the building the following year to another school. Elementary teachers in this study received many pieces of equipment and materials for which they had no storage room to secure them in. Every inch of useful space in the classroom was dedicated for student desks. Some schools converted common areas for storage; however, some supplies disappeared over time.

**Component Concept: Community engagement**

There was no policy promoting engagement of parents and/or community partnerships in this district beyond standard parent-teacher conference nights and other parent events initiated by building administrators. It was clear that barriers needed to be broken to educate parents on how to get more involved in the school and supporting their children. For the most part, parents were contacted only when their children were absent (attendance rates are part of school report cards) or if there was a disciplinary issue. Mr. Delton, a sixth-grade middle school teacher, shared his experience with parent engagement in his school during a focus group session in 2007: “The majority, I think, want the best for their children. They just don’t know how to provide support.” Mr. Garner, an elementary teacher, described the level of parental engagement at his school.

We have some parents that come into the classrooms and participate on field trips and things of that nature, but I would probably say about 20% of the parents do that. The other 80% are the ones that either have to work 18 hours a day to keep food on the table or don’t care (Interview, fall 2009).
The surrounding region in which this district is situated had a wealth of potential community partnerships, including a military base, nature center, science and technology industry, and many parks. Many schools across the nation have harnessed these resources and utilized coteaching experiences and guest speakers and partnered for funding to support learning outside of school. Teachers suspected the emphasis on test preparation had decreased the value of engaging the community in partnerships with the district. Mrs. Miken, a sixth-grade teacher said, “The pressure to cover all of our content doesn’t leave room for fun things like field trips any more. Those take money too.” Authentic learning experiences often could not fit into 20- or 45-minute segments, and this level of learning was not assessed.

The Grounded Theory: Educational Turbulence

The grounded theory emerging from this study is something we have termed as “educational turbulence.” A diagram of the concepts that comprise the theory is found in Figure 1, including interactions between the core concept federal accountability policy and the seven component concepts. Existence of and interactions between macro- and micro-level policy described as core and component concepts create educational turbulence that teachers, administrators, and students in the U.S. experience. Federally mandated school accountability has directly influenced school funding and district strategic planning—responding to the call for increased student achievement on predominantly multiple-choice, state-developed assessments in reading and mathematics through making performance gains top priority, often at any cost. Component concepts (funding, personnel, scheduling, curriculum and instruction, learning environment, accountability, and community engagement) reflect district or micro-level explicit and implicit policies that produce educational turbulence, stifling the possibility of systemic teacher quality reform. Figure 1 illustrates the direct influence of federally mandated
accountability policy on district operational funding and strategic planning, as well as the enacted plan through micro policies in the seven component concepts. Additionally, those seven component concepts interact with each other in two-way relationships.

Discussion

This study revealed the influence of macro-level educational policy on the enactment of micro-level explicit and implicit policies in a large, urban school district in the southwestern United States. The grounded theory articulated in the findings of this study, *educational turbulence*, made the full implementation of reform extremely challenging. The influence of macro-level accountability permeated every facet of micro-level policy and the district strategic plan in this study.

According to Darling-Hammond (2010) and others, teacher quality is the key to student academic success. The program in this study focused on improving teacher quality through whole-school, sustained, professional development, which produced gains in teacher effectiveness (capability) and beliefs and values regarding teaching and collaboration (culture). However, despite tremendous gains in these areas (e.g. Author et. Al, 2007; Czerniak, Beltyukova, Struble, Haney, & Lumpe, 2005; Desimone, 2009) micro-level policies in the areas of assessment, funding, personnel, scheduling, learning environment, curriculum and instruction, and community engagement created barriers to systemic reform. Educational turbulence has both macro (state and national) and micro (district and school) levels, and the interplay of components within and across levels can facilitate or impede educational change. Our findings support arguments made by Hess (1999), including his claim that existing research “suggests that
reforms fail because of inadequate implementation, planning and coordination, precisely the problems that result from policy churn” variables exacerbated by educational turbulence.

Findings in this study further extend our understanding of Fullan’s (2006) change theory. Specifically, there are two new components that educational turbulence has brought to light as needed for educational reform today, which should be considered in change theory. First, accountability policies emerged as a primary distractor issue in this study. Fullan (2006) did not include accountability issues within his fourth premise of change theory, changing the context. Accountability should be considered along with union issues, bureaucracy, and managerial issues that are currently included in change theory. Second, in the era of accountability there is an immediate need for what may be considered “quad-level” engagement. Specifically, within the sixth premise of change theory, in addition to school and community, district, and state, permeable connectivity is needed with national or federal education stakeholders. The disconnect between national policy and state implementation of policy produced the setting for educational turbulence within the large school district in this study that enacted micro-policies attempting to answer the call for increased performance on state measures of student outcomes.

Implications

Federally mandated accountability coupled with state-level control of standards and assessments have created a perfect storm to derail the ultimate goal of federal policy: educational reform. As we learned from our longitudinal study and intensive work with this district, even the most research-based and proven models of professional development will fall short of the mark if scaled up within settings where educational turbulence exists that hinders the inclusion of high-quality science teaching within the curriculum. Furthermore, alignment issues between desired federal outcomes and implemented state and local measures continue to undermine the very
educational system that is the target of national reform as evidenced in the district in this study. Undoubtedly, this is not an isolated case, as even our research team has experienced these dynamics in other districts across the U.S. Increased focus on improving performance in reading and mathematics has resulted in significantly decreased emphasis on science education for the past several years. As this study demonstrated, elementary schools rarely teach science and middle schools which do include science are burdened with trying to bridge extensive gaps in background knowledge and experiences. Turbulence in this district, and others across our nation has further left our children behind in the local and global marketplace.

Based upon experiences gained working intensively with a large, low-performing, urban district in this study some suggestions for addressing turbulence in science education reform may be warranted. At a minimum, more structured guidelines from the federal government in regard to quality and format of state developed assessments should be put in place. Ideally, common national standards (e.g. Next Generation Science Standards, Common Core) and guidelines for assessments may be better strategies to address persistent opportunity gaps in the U.S. All states should be required to teach similar content within similar grade bands, and student progress should be measured for all using high-quality measures grounded in applications to the real-world. If it is the case that educators are teaching to the test, ensuring that assessments measure student proficiency of academic standards is a place to start beyond recall of facts. Essentially, assessments should be constructed to measure student ability to apply, analyze, synthesize, and evaluate knowledge across disciplines. Critical thinking, problem solving, communication and collaboration are a few examples of 21st-century skills essential for success which can be enhanced through the teaching of science, though attainment of these skills is rarely assessed.
Our global knowledge and technological base are growing so rapidly it is no longer acceptable to place acquisition of isolated facts at the centerpiece of accountability.

As new national content standards are in the process of being developed and implemented in the U.S., it is fortuitous to align assessments with desired outcomes, including abilities and skills. It may also be wise for the U.S. to examine other systems in countries that are more academically successful to learn from their experiences and consider transformation of accountability policy. In the interim, science education reform efforts must support participants to navigate “the inevitable turbulence associated with reform” (Author, 2011a, p. xv). The most impactful, lasting outcome of delayed action will be diminished ability of systems to transform practice and educational gaps that children in the U.S. will live with for the rest of their lives—making them less competitive for employment and equally less skillful in navigating adulthood in the 21st century.

References


Author (2011a).

Author (2011b)

Author & Other (2010).

Author et al. (2010).

Author & Other (2009).


http://www.nap.edu/catalog/11463.html


http://www.michaelfullan.ca/Articles_06/06_change_theory.pdf


Table 1

Sources of Educational Turbulence

<table>
<thead>
<tr>
<th>Selective Codes</th>
<th>Axial Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mandated Accountability</td>
<td>District assessment quality &amp; format</td>
</tr>
<tr>
<td></td>
<td>Pacing guides</td>
</tr>
<tr>
<td></td>
<td>Instructional time devoted to assessments</td>
</tr>
<tr>
<td>Funding Limitations</td>
<td>Teacher salary</td>
</tr>
<tr>
<td></td>
<td>Per pupil funding</td>
</tr>
<tr>
<td></td>
<td>External grants</td>
</tr>
<tr>
<td>Mandated Curriculum &amp; Instruction</td>
<td>Mandated pedagogy</td>
</tr>
<tr>
<td></td>
<td>Accommodations</td>
</tr>
<tr>
<td></td>
<td>New curriculum</td>
</tr>
<tr>
<td></td>
<td>Class size</td>
</tr>
<tr>
<td>Personnel</td>
<td>Workload</td>
</tr>
<tr>
<td></td>
<td>Professional development</td>
</tr>
<tr>
<td></td>
<td>Job assignment</td>
</tr>
<tr>
<td>Scheduling Issues</td>
<td>Instruction</td>
</tr>
<tr>
<td></td>
<td>Collaboration</td>
</tr>
<tr>
<td></td>
<td>Preparation</td>
</tr>
<tr>
<td>Learning Environment Quality</td>
<td>Resources</td>
</tr>
<tr>
<td></td>
<td>Outside experiences</td>
</tr>
<tr>
<td></td>
<td>Facilities</td>
</tr>
<tr>
<td>Lack of Community Engagement</td>
<td>Parental involvement</td>
</tr>
<tr>
<td></td>
<td>Community partners</td>
</tr>
</tbody>
</table>