

Toward Meaningful Assessment: Lessons from Five First Grade Classrooms

laura r. kates

commentaries by

martha foote

susan freeman

marjorie siegel

www.bancollege.edu

bank street college of education

MAY 2011

occasional
paper
series

26

Toward Meaningful Assessment: Lessons from Five First-Grade Classrooms

Laura R. Kates

The assessments are simply a way for teachers to better understand exactly what a child has learned — and what he or she still needs to learn.

– Joel Klein

Rationally constructed reform strategies do not work... change (planned and otherwise) unfolds in non-linear ways... and paradoxes and contradictions abound.

– Michael Fullan

Are teachers who are faced with mandated assessments more likely or less likely to explore their students' performance in depth and use their discoveries to enrich learning?

This is the story of how six first-grade teachers in New York City responded to a mandated performance assessment—and how that response compared to a set of informal, self-selected assessments they'd identified collaboratively. This study traces the ways in which the teachers came to socially construct the mandated assessment as a bureaucratic burden rather than as the rich resource for instructional improvement its developers intended it to be. Although the assessment they were compelled to use was extremely similar to the ones they chose to use, the teachers perceived the voluntarily administered instruments to be highly pertinent to instructional decision-making, while they saw the mandated instruments as irrelevant and burdensome. The teachers' responses to the mandate call attention to the fact that the policy environments' current emphasis on scrutiny and punishment diminishes rather than enhances assessment's potential to bolster teacher learning and student achievement.

From 1998 to 2003 (and, in a highly revised form, for several years thereafter), the administration of the Early Childhood Literacy Assessment System (ECLAS) to every K-2 student was mandated throughout the New York City school system. Developed by a team of local educators, and deliberately devoid of any mechanism for reporting out, ECLAS was intended to help teachers identify children's strengths and needs in order to inform classroom instruction.

In the school I studied, ECLAS—despite its teacher-centered aims—set off a chain of interpretations and actions that ran counter to its goal of operating as a catalyst for improvement in practice. The teachers came to define ECLAS as a tool of bureaucratic control rather than as the user-friendly set of materials for informing instruction its developers intended it to be. Once ECLAS was so defined, the teachers neither analyzed it nor talked about it in ways that extended their own or their students' learning. This study bears evidence, then, to how extensively the situated meaning that is made of mandates shapes the ways in which they will be used and understood.

The profound influence of such meaning--making upon the enactment of mandates is almost always overlooked in today's policy environment, in which the cause and effect relationship between the introduction of curriculum and assessment policies and the actions teachers will take in response to them is depicted as simple and certain. For example, in a letter to parents to introduce a costly, complex on-line "interim assessment" system, New York City's former schools chancellor, Joel Klein wrote:

The assessments are simply a way for teachers to better understand exactly what a child has learned—and what he or she still needs to learn. They will help teachers identify what's working—or not working—for each child and enable teachers to better meet each child's learning needs. (Klein, June, 2003)

These comments are typical of the tenor of the talk surrounding such initiatives.

The rhetoric implies that as long as central authorities provide a standardized, compulsory means of placing information about students in teachers' hands, good outcomes will surely follow: teachers will consult the assessed information readily, and they will have the capacity, will, resources, and professional discretion necessary to translate it into relevant, effective, individualized instruction. This study casts doubt upon

the assumption that the process by which teachers formulate actions in response to mandates is so predictable, consistent, and manageable.

Studying Meaning-making About a Mandate

I conducted this research at P.S. 200,¹ a small, ethnically diverse school located in a quiet residential neighborhood in an outer borough of New York City. From mid September to mid May of the 2002 – 2003 school year, I visited the school weekly. During each visit, I observed in one or more of the first-grade classrooms during the morning literacy block and occasionally during afternoon instructional periods, too. I spent prep and lunch periods talking to the teachers and observed during staff meetings, grade conferences, literacy team meetings and staff book clubs and study groups.

Numerous interviews were conducted. The observation and interview processes were highly interrelated in that findings from observations frequently shaped the formulation of interview questions, and the focus for observations was often determined by material from interviews. I conducted multiple interviews with each of the six teachers who taught first grade as well as with the principal, assistant principal, staff developer, magnet grant coordinator, Reading Excellence Act grant coordinator and curriculum consultant. Off site, I interviewed a former curriculum consultant, the district director of early childhood, and the director of the Office of Research, Development and Dissemination at the NYC Department of Education's Division of Instructional Support.

¹ The school name and the names of all individuals quoted throughout the essay are pseudonyms.

All interviews were semi-structured. I created written protocols with open-ended questions ahead of time, but I departed from them when interviewees' responses went in unforeseen but illuminating directions. However, I always inquired into perceptions of ECLAS's purposes and possibilities. Based on an interviewee's position and history with the school and the district, I chose additional topics. These included responses to school, district and central Department of Education curriculum and assessment policies; perceptions of the sources of influence that shaped school and district approaches to curriculum and professional development; and the ways in which collaboration with colleagues impacted teaching and learning. In the fall of 2004, all six of the participating teachers were given a draft of the findings. Based on their responses, I made revisions to ensure that their perspectives were well represented in this account.

**Teachers Talk About Mandated Assessment:
“Nothing Springs Too Much Off ECLAS.”**

The first grade teachers at P.S. 200 believed that the main purpose of ECLAS—which was often referred to as a test—was the attainment of a score that could be used as the basis for the formation of guided reading groups and for the choice of a text level for those groups. Again and again, during interviews, this function was the one talked about first and foremost. When asked how they used ECLAS, the teachers consistently responded with such comments as, “I just use it to get them into groups,” “You just kind of get the levels of the children,” “To form the groups. That’s basically what I use ECLAS for. Really, honestly, I don’t use it too much more than that.”

ECLAS provided both quantitative and qualitative information about students. In addition to yielding a numerical score, ECLAS included instruments for analyzing students' meaning--making and print-decoding strategies, their fluency, their understanding of story elements and plot devices, and the kind and quality of their expressive language (see Appendix A). The P.S. 200 teachers, however, situated ECLAS almost exclusively as a source of quantitative information. For example, Regina Ferrugia, a first-grade teacher in the school for eleven years, describing her use of ECLAS results, said:

I sit down and I look at the scores, I look at what level it tells me the children are at and then I say, "Okay, this child I'll put in this group, this child I'll put in that group." —But most of the time, the ECLAS levels don't match with the guided reading levels, the children have to go up or down. You kind of put them in that level but you have to move them. So it's not very useful.

Regina's focus on the numerical dimensions of the instrument was a common one in the grade group. The guided reading levels she refers to derive from a system, developed by Fountas and Pinnell (1996) at Ohio State University and employed throughout the school and the district, in which the letters of the alphabet are used to rank student texts on a continuum of difficulty levels, with "A" being the easiest and "Z" the hardest. The texts the P.S. 200 teachers used for both guided and independent reading were organized according to this leveling system, while the ECLAS texts used a numerical leveling system. Although each ECLAS Reading Strand level was intended to correspond to a Fountas and Pinnell level, and a chart of the correlations had been provided to teachers (see Appendix B), this correlation was often seen as problematic. For example, Theresa Polito, a twenty-year veteran of P.S. 200, voiced critiques similar to

Regina's as she explained that in some cases there was no ECLAS level equivalent to a particular Fountas and Pinnell level, which therefore rendered the assessed information of little value:

I use ECLAS to put them into groups and then as we go along I fine-tune it. And if they're at a certain level, after a couple of sessions I'll realize that he's level 3, he's level 2, and I'll adjust. Generally they need about a level *B*. There's a jump from ECLAS level 3, level 3 is a *C* book and level 4A is an *F* book and level 4B is an *I*, so if I jump from *C* to *F*, what about, *D*, *E*? It's too much of a vast range and they need to get a closer range, and they need to have more books on a level because it's very limited. Another problem with ECLAS.

Theresa and Regina's remarks highlight the extent to which ECLAS was defined as a technical tool. Analysis of its results was expected to consist of a rote process of translating scores from the Reading Strand into levels for the formation of groups and the choice of texts. When they found that this process could not be carried out in a seamless, straightforward manner, the teachers believed that they had little to learn from ECLAS.

Since precision was seen as so important, grade inflation was also frequently referenced as a source of ECLAS's unreliability and irrelevance to practice. During the 1999 and 2000 school years, the district collected every school's ECLAS scores and then gave every principal showing the percentage of K-2 students meeting state benchmarks for reading in each building in the district. Although the score collection had ceased by the time I conducted this research, the first-grade teachers believed that (unlike themselves) the kindergarten teachers held on to the (now unwarranted) concern that their scores would be taken as a reflection on the quality of their teaching and thus were inclined to inflate the numbers. First-grade teacher Stella Medina's anecdote about a

colleague's encounter with the kindergarten scores is typical of grade group's stories on this subject:

One of the first-grade teachers was doing the sight words and the kid came to a level 4 in kindergarten and she went to a level 5 because that's what you're supposed to do and the kid had no clue so she went down to 4, no clue so she went down to 3, still no clue, she had to go all the way back to 2. Supposedly he's supposed to be on 4 and in her mind, she's like, "Oh well, the kindergarten teacher wanted to make me think her kids did great so she put him on a level 4."

The kindergarten scores mattered for the first-grade teachers because they were to be used to determine at which level to begin administering the assessment in the fall.

Carrie Collins, the junior member of the grade group and a co-teacher in the first-grade inclusion class, described the problems encountered in attempting to use the scores for this purpose:

At the beginning of the year, we formed our guided reading groups based on the ECLAS scores but that was a little bit of a fantasy because some of the kids the kindergarten teacher had put on a level 3 and that transfers over to what? An *F*? When we got the folders from kindergarten, this is right at the beginning, like the first two weeks, we looked at what the kindergarten teachers said their levels were. Now, level 4 they said a lot of them were at which is *I*, that's where they need to be at the end of the year. The kid was really at an *A* or a *B*, like, if you handed any of our kids an *I* book right now, they wouldn't be able to read it. And I think that's why a lot of teachers don't really take anything from ECLAS for instruction in the classroom. It's not accurate.

Here, Carrie articulates the pervasive perspective about why so little can be "taken" from ECLAS: teachers see it as an issue of accuracy. Understood strictly as a source of quantitative information, ECLAS must be numerically precise to be informative. But the teachers were required to re-administer each strand and establish a

new score for each student anyway, so if ECLAS had been defined differently, the unreliable kindergarten grades might have been experienced as a minor inconvenience rather than as a major impediment to informing instruction. However, since ECLAS was perceived as little more than a mechanism for ascertaining guided reading levels, the grade inflation—along with inaccuracies and flaws in the leveling system—became a major factor in how teachers interpreted ECLAS’s purposes and possibilities. Kathleen Harrigan, Carrie’s co-teacher in the inclusion class, articulated the pervasive perspective: “Nothing springs too much off ECLAS.”

Given their perception of ECLAS’ purposes, the teachers’ critiques made a lot of sense. Since they saw ECLAS only as a source of information for levels and scores, the position that they could obtain the same information more efficiently and more effectively using books and materials of their own choosing, was quite reasonable. As has been described, however, ECLAS also provided a body of qualitative information whose applicability to practice would have been harder to dismiss on the basis of technical shortcomings. Intriguingly, in contrast to how they analyzed ECLAS results, the teachers attended almost exclusively to the *qualitative* results of the informal assessments that they administered voluntarily. This was the case in spite of the fact that many of these assessments were nearly identical in form to ECLAS.

**Teachers Talk About Informal Assessment:
“It Really Goes Back to the Needs of the Children.”**

ECLAS’s developers in New York City relied heavily on New Zealand literacy expert Marie Clay’s Observation Survey for the design of their own instrument. The most

widely known and used tool in Clay's survey is the Running Record. The ECLAS Reading Record (see Appendix C1) was a simplified version of Clay's Running Record (See Appendix D). To take a Reading or a Running Record, the teacher sits beside a child and uses a standardized set of notations to make a record of what the child says and does while reading aloud a portion of text. Along with providing an accuracy rate, the Reading or Running Record provides teachers qualitative information about the student's meaning making. By listening to the child talk about a book and by analyzing the nature of her or his errors, the teacher learns which decoding and comprehension skills and strategies the child is independently applying and which need to be supported with further instruction.

By 2002 all of P.S. 200's first-grade teachers had been trained in how to take Running Records and administered them regularly in their classrooms. When I sat down with the teachers to talk about how they used their Running Records, the tenor of their talk was very different from the technical tone of the discourse about ECLAS. According to Theresa, for example:

It really goes back to the needs of the children ... I'll take notes and I'll know what to focus on, things that the kids had trouble with, things that they had in common that they're having trouble with, what they're doing well, solving words, comprehension, strategies that they're using or not, I'll bring that up. Years ago, before we had this, it used to be a blur.

In contrast to her critique of ECLAS' leveling system, Theresa expects her Running Records to feed and further her professional decision-making, not to substitute for it. Stella, a member of the staff since 1989, is also much less concerned with strict adherence to a particular system, as indicated in the example she provides about how her Running Records inform her planning:

I was doing the Blue Group. A couple of them got stuck on the word “tiny” ... They were saying “tight” instead of “tiny,” they were looking at the picture clues which were things about shoes and how the shoe would maybe be tight. So a strategy that I’ll work on with them is to look at the whole word, don’t just look at the beginning because the initial consonant “t” is what they were looking at and the picture cue is what they were looking at ...

Here, Stella uses her Running Records to engage in a non-standardized decision-making process about her students’ needs. She melds her insights into their reading behaviors with her knowledge of emergent reading strategies to identify an instructional course of action. Although ECLAS also provided just the sort of information Stella describes, about which strategies were being used or not (see Appendix C2), the first-grade teachers did not look to it for such purposes. While they defined the ECLAS record as an instrument whose only purpose was to provide them with a reliable formula for leveling and grouping, they defined their Running Records—which in design were almost identical to ECLAS—as a resource for expanding the insights and information upon which they could base their professional judgments.

Of course the question arises as to why the teachers *would* consult ECLAS when they knew how to obtain the same information from other sources. It needs to be borne in mind, however, that they were required to administer ECLAS, so its results were right in their hands. Additionally, they frequently mentioned how challenging it was to find the time to administer their own informal assessments during the ECLAS administration periods, which lasted about six weeks both in the fall and in the spring. The question then becomes one of why—since they had the ECLAS information and the teaching structures that made its results highly pertinent to practice—they chose to ignore the ECLAS results

while seeking the same sort of insights from almost identical instruments, at great inconvenience.

It was in considering this question that I came to believe that the teachers' responses to ECLAS had as much to do with how the instrument was socially and institutionally *defined* as with how it was actually *designed*. In particular, I found that the school and the district played prominent roles in ECLAS's coming to be situated as an unwieldy technical tool rather than as the rich resource for instructional improvement its developers intended it to be.

Stream of Influence: District Score Collection

Like the teachers, the district leadership focused almost exclusively on ECLAS's quantitative dimensions. During the 1999 and 2000 school years, each elementary school principal was required to submit their ECLAS scores to the district office for analysis and comparison. Anna Franco, the district early childhood director, described the decision to take this action as rooted in both a concern that without external monitoring, the assessment might not be administered fully or correctly in some buildings and in a belief that making the scores public might act as a lever for improved instruction, especially in the kindergartens (about which she had grave concerns):

I knew there was trouble in kindergarten. There was not much teaching going on. It was seen as a grade where when you got up in the morning and got in you'd decide what to do. So my purpose in doing this was just to let teachers know that somebody was looking at it, so make sure you do it and the next reason was to see where there were trouble spots.

As discussed above, at the district office, the ECLAS Reading Strand scores were

correlated with state benchmarks for reading and a grid was created showing the percentage of K-2 students meeting the benchmarks in each school. This grid was then distributed to every elementary school principal during a principal's conference, which made each school's standing in relation to the others highly visible to everyone.

According to Anna, for a limited time this approach did have some of the benefits she sought. After a while, though, unanticipated disadvantages surfaced:

When principals see things like this published, they start putting a lot of pressure on the teachers because they say, "Hey, you're making me look bad." And unfortunately, many principals don't understand kindergarten so one principal told his teachers, "Well I'm not ordering any paint anymore because you're not gonna be painting." So you always have to watch for doing something like this.

In the fall of 2001, after hearing teachers' concerns about the score collection's detrimental effects, Anna decided to retract the policy. Her sincere efforts to be responsive to the teachers and to shift the emphasis back to informing instruction did help many people, as Theresa Polito put it, "... feel more relaxed about the test." However, as a result, the perception of ECLAS as an assessment primarily geared toward external evaluation and accountability, rather than toward teachers' own decision making, was also cemented in the minds of many. Curriculum consultant Emily Roberts described the impact of the policy this way: "The teachers don't look back at it to see where it could be helpful because it was imposed with that stigma attached, so [they ask] - 'How could it be directly what I would need? How could it be individualized enough for my particular needs?' The score collection, then, was one of the factors that fueled the teachers' construction of ECLAS as an instrument ill suited to informing personalized decision making about students' needs.

By requiring submission of scores, by making the scores public, by taking the scores to be adequate indicators of the quality of practice, and by creating a climate in which the scores were linked in practitioners' minds to the possibility of incurring undeserved sanctions, the district sent the message that ECLAS was a tool of bureaucratic control rather than a resource for teacher learning. Having come to define ECLAS as a means of monitoring student success and teacher efficacy from afar, the teachers were disinclined to include it in the pool of professional resources they drew upon to fuel the inquisitive, collaborative learning processes they undertook in order to improve their practice. In their authoritarian, if well-intentioned, attempts to use ECLAS as a lever for lifting the quality of what took place in the classrooms, the district leaders inadvertently diminished rather than deepened ECLAS's potential influence on instruction.

Stream of Influence: The Design of Learning Opportunities

The design of opportunities to learn about ECLAS also emphasized the procedural and the prescriptive over the subtle and the substantive. At P.S. 200, teachers described two very different types of professional development offerings. Some, which teachers welcomed and found helpful were open ended and responsive to individual teachers' experiences, interests, and priorities. Such opportunities included demonstration lessons and grade meetings facilitated by curriculum consultants, informal conversations during shared preparation periods, and content-focused study groups held by the district office. Teachers contrasted such enriching activities with events that had overly rigid, pre-set agendas and that emphasized external accountability over genuine inquiry. For

example, in connection with her participation in a professional development course, Theresa was required by the district to implement a teachers' book club. Lydia Mendez recounted the grade group's experiences with that club. As she described it, since it was something no one especially wanted to participate in, the group selected a book they had read already:

It's nonsense you know really, "Pick a teacher book. Have a discussion among the first-grade grade teachers." ... All this stuff is like make-believe ... It's too bad and you know why it turned out to be this way because it's so much pressure that everything becomes meaningless. I mean it's never like, "If you'd *like* to, if you *want* to, if you *need* to." If you *have to*, it takes away the joy.

Although the teachers liked to engage with new materials and ideas, they did not think that instances of "contrived collegiality" (Hargreaves, 1994, p.17), such as the book club, provided the opportunity to undertake the kind of in-depth, reflective, interactive processes that they believed made for genuine learning. When they encountered ideas about practice in this kind of coercive context, they were, understandably, disinclined to thoughtfully engage with the subject matter and instead often turned to the sort of cynical, superficial responses Lydia described.

Such cynicism was often in evidence during professional development events about ECLAS, which, unfortunately, were invariably contrived and controlled. Those events were facilitated by the school's ECLAS coordinator, Mary Dowling. Each school in the district had been required to appoint such a coordinator, who would attend district meetings and then disseminate the information received at the district office to the school staff. There weren't many first-grade meetings about ECLAS and their purpose was always for Mary to share new required procedures for scoring or record keeping.

One November morning, for example, Mary began an ECLAS meeting by announcing, “There’s no good way to say this, so I’m just going to say it. There’s a new phonemic awareness component to ECLAS,” to which the usually reticent Stella Medina asked, “If you finished your fall ECLAS, you don’t have to do it?” The group snickered as the always straight-talking Lydia added sarcastically, “We do so much assessment, when do we teach?” As Mary tried to talk through the details of the new component’s purposes and procedures, the teachers continuously interrupted her to air gripes about the logistics of administering it. At the meeting’s conclusion, when Mary offered a handout about upcoming professional development events related to ECLAS that were being conducted by the Department of Education, it was left on the table untouched as the teachers filed out of the room.

On another occasion, Mary brought the first-grade group together to show them some new recording forms that teachers in a neighboring district had created for ECLAS results. Regina interrupted Mary’s presentation to announce, “It’s very difficult to do this at the beginning of the year. Everything gets thrown out the window. You have to say to yourself, ‘My time is wasted. My routines are shot.’” The other teachers nodded emphatically as Mary offered to ask the principal to send more substitutes to cover the classes while ECLAS was being administered. This suggestion offered little satisfaction, however, as Regina responded, “I find when the sub comes in, you have to take care of *them*. Last year, I went out in the hall with a kid, when I came back in the sub was sitting there by herself reading a big book while the kids ran around the room.” As the teachers giggled, Mary replied, “I can’t vouch for the quality of the people you’ll be getting, but at

least you'll have another body in the room." Theresa irritably responded, "What do we need all these assessments for? We know what we're doing."

These excerpts are typical of the tenor of the talk about ECLAS, which was dominated by expressions of resentment about the practical demands of administering the assessment and of resistance to having to administer it at all. If the school and district leadership had used the meetings to focus on the more substantive aspects of ECLAS procedures and purposes, the teachers' discourse might well have been more substantive, too. The quality of conversations during professional development events that emphasized genuine inquiry into meaningful issues of practice, was very different. For example, during the year that this research was conducted, the first-grade teachers were collaborating with the literacy consultant Emily Roberts on rethinking their spelling curriculum. During her first meeting with the first grade the previous week, Emily asked what they would be interested in focusing on together in connection with the district-wide spelling initiative. There was some discussion of the difference between studying sound patterns and studying visual patterns in words. The first-grade then gathered, during the meeting excerpted below, to reflect on a demonstration lesson that Emily had conducted that morning in Kathleen and Carrie's classroom:

Theresa I'd like to go back to the difference we talked about last week between searching for spelling patterns and searching for sounds because Kathleen and I were talking and we were a little bit confused. So, with the spelling patterns, you're looking for groups of the same letters together but they won't necessarily have the same sound?

Emily That's right.

Kathleen Why would you use the spelling patterns and not the sound?

- Emily** Because developmentally it's easier to find spelling patterns than to sort by sound.
- Kathleen** I tend to go more towards the sounds.
- Lydia** Me, too.
- Emily** They need to know that spelling patterns have different sounds.
- Kathleen** So if you were doing /ache/, you'd say look for a-k-e?
- Theresa** It's really like that when these letters are together in these words or this sentence, they make this sound.
- Regina** It's like at the beginning of the year when we do letter searches, I say, 'Look for a letter, not a sound.' So, in other words, you should really stick more like to patterns than to sounds?
- Emily** Usually they'll be thinking about it visually, then the auditory comes next.
- Theresa** [Nodding.] Initially. Initially you're doing it visually.

During this interchange, the teachers showed none of the preoccupation with the mundane logistics of implementation that characterized their conversation about ECLAS. Instead, they probed the philosophical rationale underlying the suggested practices, analyzed the efficacy and relevance of past practices and questioned the facilitator and one another in order to begin to construct a precise mental picture of the particular form these activities and procedures might take in their individual classrooms. Because the agenda for this meeting was shaped from the outset by the teachers' own definition of a worthwhile focus, they demonstrated a high degree of comfort with participation in a substantive exchange of ideas.

Similarly, when the teachers and the administration decided together that some of their weekly grade meetings ought to be used to refine the design of the teachers' own

informal assessment systems, the focus remained squarely on implications for teaching and learning. During these sessions, the teachers shared the binders, folders, and notebook systems that they had created for assessing children during guided and independent reading. They discussed which organizational approaches were working well and which needed re-thinking. All of the teachers cited this cycle of meetings as an example of a useful, welcome professional learning event. Theresa summed up the group's positive experience:

We asked each other, "What are you doing this year? How are you going to keep notes? What are you going to do for record-keeping?" We really wanted to make it work. We really wanted to get it right. The system I have now keeps me on target; it makes me more effective. Teachers talking with teachers often generates good ideas.

There were never meetings like these, (i.e., with flexible agendas and a focus stemming from teachers' self-selected questions or areas of inquiry) about ECLAS. Instead, like the score collection, the approach to professional development about ECLAS sent messages about its purposes that were antithetical to the teachers conceiving of it as an opportunity to learn. The emphasis on the routine, mandated aspects of administering ECLAS furthered the sense that it was not really designed to serve individual instructional purposes, as was claimed, but was instead actually geared toward meeting the externally conceived, impositional, often irrational, objectives of higher-ups.

The school and the district, then, operated as key streams of influence in shaping teachers' socially constructed beliefs and practices about ECLAS. Despite the sound instructional intents that seemed to form the basis of decision making at all levels of leadership, what was bureaucratically issued came to be bureaucratically defined. What

was bureaucratically defined, in turn, was not simultaneously situated as a rich resource for professional learning and improvement of practice. This study highlights, then, both how the very profound impact that site based meaning making has on how a mandated assessment will be used and understood and also the key role that social organizations that surround teachers play in shaping such meaning making.

Implications for Practice

As they engaged in the daily challenge of enacting effective instruction for their diverse groups of six- and seven-year-olds, the P.S. 200 teachers met up constantly with mandates and demands from the central Department of Education, the district and the school. Some of these mandates and demands were experienced as generating exciting opportunities to expand the teachers' ever-developing capacities; others, like ECLAS, were experienced as constricting those capacities in frustrating and needless ways. The teachers' story, then, embeds important lessons for the crafting of policies and professional development likely to nurture substantive sense-making processes and limit impediments to learning.

As discussed previously, the district fueled teachers' positioning of ECLAS as a flawed, unreliable technical tool by collecting the ECLAS scores and publicizing among principals the percentage of students meeting benchmark reading standards in each school. Nonetheless, professional development events that encourage teachers and principals to take an inquiring stance toward the results of mandated assessments might help move the discourse about such assessments away from the trivial and the technical. For example, rather than being subject to such mundane activities as examining grids and

learning procedures, principals and teachers might be given release time to examine a mandated instrument carefully—something they will rarely have time for in the course of an ordinary work day. Teachers also need the opportunity to clarify confusions about the meaning of particular aspects of the assessment and talk together about the implications of the assessments' content for their school's approach to curriculum and instruction.

For those leaders bold enough to make them so, professional gatherings could also be used as forums in which to critically consider the meaning of quantitative data. Rather than perpetuating the popular perspective that scores alone tell a complete story, such meetings could support principals and teachers in the process of collectively unpacking what different numbers mean in different contexts. For example, colleagues might present portraits of students who have the same scores but very different histories, experiences, and learning needs. Such presentations could call attention to the limited value of the numbers alone in coming to fully understand the instructional implications of assessment results.

Additionally, school and district leaders can send a more substantive message about a mandate's meaning by involving teachers in decision making with respect to local compliance. In my extensive experience as a curriculum consultant in numerous New York City schools, I found that even very collaborative schools tended to shut teachers out of discussions about the design of local responses to centralized mandates. As a result, school policies are often made that are in tension with teachers' pre-existing inquiries, agendas, schedules, and routines, heightening the teachers' experience of the reforms as irrational and impositional. If teachers had more say over the particular form

that school-level compliance ought to take, along with greater opportunities for in-depth exploration of an assessment's content, the impulse to attend exclusively to the mandate's most obvious and easily understood dimensions might not be as prominent as it was at P.S. 200.

Implications for Policy

When ECLAS was adopted in 1998, New York City's central board chose a reform strategy in which a mandated performance assessment was to operate as the central engine of a large-scale effort to effect change in practice. The findings of this study indicate that such a strategy has some strengths, but also substantial limitations – the idea that a mandated instrument can spur a commitment to in-depth learning about new practices or to the challenging process of adopting new structures is not borne out.

In his analysis of the features of effective and ineffectual reforms, Fullan (1999) points out that, “Rationally constructed reform strategies do not work” (p.3). Drawing on the tenets of complexity theory, he attributes this to the fact that “... the link between cause and effect is difficult to trace ... change (planned and otherwise) unfolds in non-linear ways ... and paradoxes and contradictions abound...” (p. 4). The central board's strategy in no way integrated an acknowledgement of these challenging dimensions of change. Instead, a predictable, manageable process was envisioned, in which the assessment itself would both trigger insight into the need for individualized instruction and the desire to learn how to enact it. In reality, though, specifically dictating the format, materials, record-keeping procedures, and time frames for administration, opened the

instrument up to its being constructed along lines incompatible with the foundational goal of exercising influence over professional decision-making.

Counter to the intents stated by its authors, the design of the ECLAS reform invited the sorts of punitive misuses the district initially adopted. While the decision making of those working out of both the district office and the office of the Division of Instructional Support seemed to be guided by a genuine desire to raise the quality of teaching and learning, the specificity of the mandate left it highly vulnerable to being used as a tool for monitoring and evaluation which, as this study has demonstrated, in turn detracted from the policy's potential to spur improved practice.

This story of how uniformity and control sapped the capacity of a policy to function as a catalyst for teacher learning and change holds important lessons for contemporary approaches to policymaking. In comparison to so many policies currently being constructed, especially those emanating from the Federal No Child Left Behind Act, ECLAS was extremely mild in its mechanisms for monitoring and in its linkage to scrutiny and evaluation of teachers and principals. In order to emphasize informing teachers' own practice, no formal procedure for reporting out was built into ECLAS, nor was any system for issuing rewards and sanctions attached. Yet, even in this context of the deliberate downplaying of external intrusions, the impersonal, inflexible elements of the policy adversely impacted its potential to teach teachers. Such an outcome is all the more likely in today's climate, in which evaluative and punitive policy dimensions are deliberately highlighted as self-evident guarantors of effective change. The findings of this study indicate that this is precisely the wrong paradigm in which to root a reform if

the self-proclaimed goals of bolstering teacher skill and student achievement are to be meaningfully borne out.

References

Fountas, I., & Pinnell, G. (1996). *Guided reading: Good first teaching for all children.*

Portsmouth, NH: Heinemann.

Fullan, M. (1999). *Change forces: The sequel.* London: Falmer Press.

Hargreaves, A. (1994). *Changing teachers, changing times.* New York:

Teachers College Press.

Appendix A: Overview of ECLAS Strands

Overview of Literacy Development Checklist

STRAND	Level 1 Getting Ready to Read and Write	Level 2 Emergent Reader/Writer	Level 3 Beginning Reader/Writer	Level 4 Developing Reader/Writer	Level 5 Semi-Independent Reader/Writer	Level 6 Independent Reader/Writer
ALPHABET/SIGHT WORDS	ALPHABET SHEET 美 • 20 letters by name (of 54)	ALPHABET SHEET 美 • 40 letters by name (of 54) • 13 letters by sound	BASIC SIGHT WORDS 美 • any 20 words	BASIC SIGHT WORDS 美 • any 35 words	ADVANCED SIGHT WORDS 美 • any 20 words	ADVANCED SIGHT WORDS 美 • any 35 words
PHONEMIC AWARENESS <i>Prerequisite: Level 2 or higher of Alphabet/Sight Words Strand</i>	PHONEMIC AWARENESS ACTIVITY 美 • no recognizable responses	PHONEMIC AWARENESS ACTIVITY 美 • most responses at Level 2	PHONEMIC AWARENESS ACTIVITY 美 • most responses at Level 3	PHONEMIC AWARENESS ACTIVITY 美 • most responses at Level 4	PHONEMIC AWARENESS ACTIVITY 美 • most responses at Level 5	PHONEMIC AWARENESS ACTIVITY 美 • most responses at Level 6
READING <i>All children, once only: Pete's Chicken 美</i> • Listening Comprehension • Reading Placement	INTRODUCTION TO PRINT 美 <i>Things I Like to Do</i> • Print awareness ENVIRONMENTAL PRINT BOOK 美 <i>Signs and Things</i> • Recognizes 8 signs/objects	READING RECORD 美 <i>Miss People's Pets</i> or <i>I Paint</i> READING ACTIVITY • 85% accuracy Writing/Discussing ACTIVITY • Comprehension (Expressive Language)	READING RECORD 美 <i>My Shadow</i> or <i>Things I Like to Do</i> READING ACTIVITY • 85% accuracy Writing/Discussing ACTIVITY • Comprehension (Expressive Language)	READING RECORD 美 <i>Nowhere and Nothing</i> or <i>Baby Bear's Present</i> READING ACTIVITY • 85% accuracy Writing/Discussing ACTIVITY • Comprehension (Expressive Language)	READING RECORD 美 <i>Last!</i> or <i>The Wild Woods</i> READING ACTIVITY • 85% accuracy Writing/Discussing ACTIVITY • Comprehension (Expressive Language)	READING RECORD 美 <i>A Weekend with Wendell</i> or <i>The Dandelion Seed</i> READING ACTIVITY • 85% accuracy Writing/Discussing ACTIVITY • Comprehension (Expressive Language)
WRITING MECHANICS	WRITING SAMPLE 美 <i>(from any Reading Record)</i> • drawing with recognizable details • random letters or letter strings	WRITING SAMPLE 美 <i>(from any Reading Record)</i> • drawing with recognizable details • relevant letters/letter strings that relate to the story	WRITING SAMPLE 美 <i>(from any Reading Record)</i> • drawing with recognizable details • words and phrases that relate to the story • beginning of correct directional pattern • phonetic/inventive spelling	WRITING SAMPLE 美 <i>(from any Reading Record)</i> • relies more on writing than drawing • at least one simple sentence • correct directional pattern • beginning use of initial capitals, spacing between words, and ending punctuation • phonetic/inventive spelling	WRITING SAMPLE 美 <i>(from any Reading Record)</i> • two or more sentences • generally correct use of initial capitals, spacing between words, and ending punctuation • beginning use of conventional spelling of age-appropriate words; phonetic/inventive spelling acceptable	WRITING SAMPLE 美 <i>(from any Reading Record)</i> • three or more sentences • essentially correct use of initial capitals, spacing between words, and ending punctuation • spelling of age-appropriate words is more conventional than inventive; phonetic/inventive spelling acceptable

Note: 美 = group activity 美 = individual activity

Appendix B: Leveling Correlation Grid

Attachment # 7 ³²
32

LEVELING CORRELATION							
GRADE LEVEL	Reading Stage	Fountas & Pinnell	ECLAS Level	Wright Group	Reading Recovery	D R A	
K & 1 ST GRADE	EARLY EMERGENT	A	2	A	1	A - 1	
		B	2	B	2	2	
		C	3	C	3	2	
	UPPER EMERGENT	C	3	D	4	3 - 4	
		D	3	E	5 - 6	5 - 6	
		E	3	F	7 - 8	7 - 8	
		F	4	G	9 - 10	9 - 10	
		G	4	H	11 - 12	11 - 12	
		H	4	I	13 - 14	13 - 14	
	2 ND GRADE	EARLY FLUENCY	I	4	J	15 - 17	16
			J	5	K	18	18
J			5	L	19	20	
K			5	M	20 - 22	22 - 24	
L			6	N	23 - 25	25 - 27	
3 RD GRADE	FLUENCY	M	6	N	25 - 27	28	
		N	6 +	O	28	30	
N			P	30	34		
O			Q		38		
P			R		38		
4 TH GRADE		Q		S		40	
		R		T		40	
5 TH GRADE		S				44	
		T				44	
		U				44	

These level correlations are suggestions only. Suitable texts are those which best match children's needs and their experiences / knowledge of the subject matter in the text - ie with the appropriate levels of support & challenge.

Appendix C1: Example of ECLAS Reading Record Materials

Sample notation: Reading Record on Script

Name Maria Miller Birth Date 2-27-91
 Teacher Jerri Davis Date 11-25-98

READING RECORD

Level 3 Book: *My Shadow* by Jean Bennett (38 words)

👁️ *Teacher reads the title and "My shadow follows me."*

My shadow follows me around the corner. ^{SC} ^{T ✓} ^{SC}

My shadow follows me over the bridge. ^{SC} ^{building ✓}

My shadow follows me down the sidewalk. ^{SC}

My shadow follows me up the steps. ^{and down ✓}

My shadow follows me through the door. ^{T ✓} ^{SC}

Look! No shadow.

89% Accuracy

4 errors
 5 self-correx
 1 repetition

# Errors	Acc. Rate
1	97
2	95
3	92
4	89
5	87
6	84
7	82
8	79
9	76
10	74

NYC ECLAS

Acknowledgment CTB is indebted to the following for use of: Text from *My Shadow* by Jean Bennett, copyright © 1996 by Jean Bennett. Reprinted by permission of The Wright Group, Bethel, Washington.
69881

Appendix D: Example of a Running Record

The following is a running record that shows a reader using a whole range of strategies. This is a record of Rory reading *Wibble Wobble Albatross*: (Ready to read, level H, Fountas & Pinnell, 1996)

Wibble Wobble Albatross	Sources of Information			
	E	SC	Error	SC
<p>W ✓ ✓ ✓ l leaves ✓ ✓ When an albatross learns to fly, ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ it leaves the land and goes out to sea.</p>	1		(M) (S) (V)	
<p>✓ li likes SC ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ It lives there for a long, long time.</p>		1	M (S) (V)	M (S) (V)
<p>✓ gl A ✓ ✓ ✓ It glides T on the wind</p>	1		M S V	
<p>✓ ✓ ✓ ✓ ✓ It rests on the water</p>	1		M S V	
<p>✓ ✓ ✓ ✓ ✓ ✓ It looks for fish to eat.</p>				
<p>✓ ✓ ✓ Now, the albatross</p>				
<p>✓ ✓ ✓ ✓ ✓ comes back to its home</p>				
<p>✓ ✓ ✓ on the land</p>				
<p>His sc his R sc ✓ ✓ R it sc ✓ It has been out at sea</p>	1 1 1		M S V M S (V) M S (V)	M S (V) M (S) (V) M (S) (V)
<p>✓ ✓ ✓ ✓ for nearly three years.</p>				
<p>Now sc ✓ ✓ walks R3 a No wonder it wobbles about T</p>	1 1	1	M S (V) M S V	M S (V)
<p>W ✓ ✓ ✓ ✓ ✓ a ✓ When it tries to walk again</p>				
<p>✓ ✓ Wibble wobble</p>				

Here Rory said leaves for learns. This error makes sense, is structurally correct and the beginning looks right.

When analyzing a self-correction, you look at the sources of information used to make the prediction, then look at what made Rory change his mind.

Rory said "it" for "at". Up to this point it does not make sense, nor is it structurally correct. It does, however, look similar. He used a visual cue to make the prediction. It looks as though he changed his mind because it did not make sense and it was not structurally correct.

