"I think I learn a lot more about my kids"

Understanding how class size reduction & assessment shape education experiences

Elizabeth Graue

Denise Oen

Erica Rauscher

Wisconsin Center for Education Research

University of Wisconsin Madison

ABSTRACT

Both class size reduction and assessment can improve of instructional quality and student outcomes. In this paper we consider how these two resources come together in Wisconsin's Student Achievement Guarantee in Education (SAGE) program, a policy widely viewed as a class size reduction initiative. Through a qualitative study in nine SAGE schools, we examine assessment practices with attention to accountability, alignment, audience, and action.   We then provide a focused examination of assessment in three schools thought to have high levels of SAGE implementation. We argue that assessment can be a synergistic system when it: 1) is accountable to all stakeholders, 2) is aligned with other elements of the instructional system, 3) considers the needs of the audiences that each assessment addresses and 4) serves as a catalyst for action that improves student learning. This work suggests the following steps in our future evaluation work:

- To most adequately address the practice-achievement link, we need to have classroom level data for sampling and analysis.

- We need to examine carefully the answers on question 9 of the teacher survey that addresses assessment opportunities and relate those answers to our case study knowledge in the nine-school sample and the instructional quality ratings on the CLASS to see how they provided convergent or divergent portraits of assessment practice.

- We also need to examine the potential of the End of Year Report to provide information on the alignment of standards, curriculum, instruction, assessments, and reporting formats.

If class size reduction changes opportunities to teach and learn, they are not irresistible; students

and teachers must use the opportunities, and that requires will and knowledge (Cohen,

Raudenbush, & Ball, 2003).

Class size reduction has been advocated as a tool for enhancing student achievement by

increasing the opportunities that teachers and students have to interact around relevant content, reducing

disciplinary disruptions, and enriching teacher knowledge of students' strengths and weaknesses (Biddle

& Berliner, 2002). The literature on class size reduction has generally supported these suppositions, but

there is some disagreement about the degree to which the considerable investment merits the costs

(Hanushek, 1999; Harris, 2006).

Better teacher knowledge of student needs is a key element in class size reduction's mechanism.

With fewer students, teachers are thought to have more intense interactions with individuals, generating

assessment information that can be used to tailor instruction to particular students. Previous research

hints at this, suggesting that students are more likely to interact with teachers (NICHD Early Child Care

Research Network, 2004) and are more focused on activities in classes with smaller groups (Blatchford,

Moriarty, Edmonds, & Martin, 2002). Researchers have not examined explicitly assessment's function in

making the most of class size reduction. With assessment playing a central role in education reform, it is

important to improve our understanding of the part it plays in teacher decision-making and action.

In this paper, we examine how two high interest resources, assessment practice and class size

reduction, come together to produce opportunities for teaching and learning in the primary grades. Based

on ongoing fieldwork in schools participating in the SAGE program, we explore how smaller groups and

assessments generate knowledge of students and instruction calibrated to student needs. We argue that

these resources, in and of themselves, do not guarantee better instruction or improved achievement.

Instead, their contributions require specific activation in particular local contexts, matching the history

and needs in a school and classroom.

Literature Review

Class size reduction has been implemented and studied in a number of contexts and through a range of strategies. It has been found to have positive effects on student achievement (Biddle & Berliner, 2002; Finn & Achilles, 1990; Glass & Smith, 1979; Grissmer, 1999; Smith, Molnar, & Zahorik, 2003) and student and teacher attitudes (Smith & Glass, 1980; Zahorik, et al, 2003). The effects of small classes seem to be the most positive in the early grades and for African-American students and students living in poverty (Biddle & Berliner, 2002; Finn, Gerber, Achilles, & Boyd-Zaharias, 2001; Smith et al., 2003) and appear to persist beyond the primary grades (Ehrenberg, Brewer, Gamoran, & Willms, 2001; Finn, et al, 2001; Nye, Hedges, & Konstantopoulos, 2001; Blatchford, 2003).

There is disagreement about how instruction is changed in smaller classes. Some researchers assert that teachers use essentially the same strategies in large and small groups (Cahen, Filby, McCutcheon, & Kyle, 1983; Rice, 1999; Slavin, 1989; Stasz, 2002). Others have been able to document subtle but important differences in teacher practice. Zahorik et al (2003) found that effective teaching in small classes was characterized by less time on discipline, clear academic and behavioral expectations, balanced instructional methods and higher degrees of individualization.

Blatchford's (2003) study of British implementation found that class size reduction enhanced learning through:

- Teacher task time with pupils
- Teacher/individual support for learning
- Classroom management and control (p. 150)

Blatchford suggests that smaller classes create the opportunities for what he calls *teacher support for learning* through effective 'scaffolding' of student activity. Assessment is certainly a key aspect of this kind of support as it is difficult for teachers to scaffold student learning in the absence of information about student needs.

Assessment has taken on increased importance in recent years, as standards-based educational reform has heightened attention to the role that evidence of student learning can play in improving

education. The promise of standards-based reform centers on the idea that standards that articulate what

students should know and be able to do provide a clarity of purpose for education systems (Darling-

Hammond, 2004). This reform has been motivated by different logics of practice. The first asserts that

standards will promote investments in schools and curriculum changes through the alignment of goals and

practice. If school people show that they have taken a systematic approach to instructional design and

have implemented rigorous measures of student learning, the public will feel more comfortable

supporting their work. The second suggests that sanctions attached to performance in relation to

standards are the engine that drives the reform. Poor performance on tests will prompt better teaching and

therefore, stronger achievement.

These two logics have worked in tension in the last ten years as schools have enacted standards-

based reform. Assessment is a central component of standards-based reform, informing instructional

practice and providing evidence of instructional efficacy. As assessment has taken center stage in current

reform practice, there has been concern about the tensions inherent in its use as both an evaluative and

instructional tool. The political power of the accountability function, with its testing systems and learning

targets, has produced a context in which "the standards movement has been corrupted, in many instances,

into a heavy-handed system of rewards and punishments without the capacity building and professional

development originally proposed as part of the vision." (Shepard, 2000, p. 9).

Shepard (2005) makes an important distinction between *formative assessments*, which take place

during instruction to inform teaching and learning and *benchmark or interim assessments* that are used to

monitor progress toward state standards and national requirements like NCLB. While the benchmark

assessments might give information about individual students' knowledge they lack the specificity or the

feedback functions of formative assessments. This distinction is relevant to our case because the utility of

assessment to improve teaching and learning rests on the degree to which formative assessments provide

information for fine-tuning instruction to specific students in specific contexts.

When we think of class size reduction and assessment as resources for schooling, it is clear that

their potential is linked to their activation in a specific context. Smaller classes *can* change what teachers

do but it is not automatic. Assessment *can* provide information to inform instruction but it can also be

highly politicized and unconnected to classroom practice. It is the specific synergy that using assessment

practice in a class size reduction context that we explore in this paper.

Methods

This paper comes out of a multiyear evaluation of the *Student Achievement Guarantee in*

Education (SAGE) program. While SAGE is primarily seen as a class size reduction program, it has 4

major requirements: reduce class sizes to 15:1, increase the number of hours the school building

available to the community, rigorous curriculum, and professional development for staff. The focus of

this report, class size reduction and assessment, addresses two of these pillars: class size reduction and

rigorous curriculum. Both elements were targets of the initial grant for a qualitative evaluation begun in

2004-5, addressed by the following research questions

- How is instructional practice related to content, grade level, and special services?

    - What is the nature of instructional practice in various configurations of class size reduction?

- How is assessment related to instruction?

    - What is the nature of assessment practice and the knowledge it generates?

    - How does curriculum alignment—with local standards, relevant district and state assessments, and
      research-based practices—mediate teacher adaptation of curriculum for individual children?

These questions were addressed initially in an integrated report of SAGE instructional and

administrative practice (Graue, Hatch, Rao & Oen, 2005) and were deemed important enough to continue

in the IQ^2 research project begun in 2005. Our interpretations of assessment in SAGE classrooms comes

out of the current accountability focus in U.S. schools and the attention paid to assessment-driven

instruction that focuses teaching on student needs. It also comes out of a focus in the literature on class

size reduction and in our fieldwork that linked the power of smaller classes to teachers knowing their

students better. Finally, it emerged from earlier work on the project (see Graue et al, 2005) that

recognized the importance of systemic analysis of teaching, learning, and administrative practice in

SAGE schools. Rather than focusing solely on teacher practice, we found that it was more fruitful to

examine how teacher practice fits into systems in schools and districts that have inter-related parts –

instruction is related to assessment, which is related to decision-making, which is related to learning.

Our research examines SAGE implementation in a sample of schools representing a range of

student achievement, urban, rural or semi-urban location, student poverty and different approaches to

reducing the ratio of students to teachers. A number of school characteristics are depicted in Table 1. The

data in this report span three school years: 2004/5, 2005/6, 2006/7. Sampling for the initial study was

based on data from 2003. Here, we present information on the schools from the most recent school year,

2006-7. Examining the characteristics, it is important to recognize that labels of high achieving and low

achieving are relative terms – schools are high achieving relative to their predicted achievement

(compared to schools with similar risk factors and resources). In addition, school achievement varies

from year to year, with most of the means based on relatively small samples of students taking the WKCE

each year. Finally, the data on which we base achievement is school level and the data on which we base

the bulk of our practice analysis is classroom level. This results in a particular kind of sampling-results

gap. The school level data represents the accumulated average effects of instructional practices across

multiple years of education, including SAGE. For example the third grade reading scores reflect efforts

made by at least 4 classroom teachers per child (K-3), averaged across all children who took the test. The

effects of SAGE are joined with the effects of other educational interventions and the effects of other

resources such as home environment and experience.

School	Bethany[1]	West Canton	McMahon	Earhart	Calloway	Montford	Allerton-Farwell	Wellstone Blvd.	Gallows
Achievement	High	High	Improving	Improving	Improving	Improving	Low	Low	Low
Geography	Urban	Rural	Semi-urban	Semi-urban	Urban	Rural	Rural	Urban	Urban
District	Mallard	West Canton	Bellamy	Maxwell	Mallard	Walton River	Allerton-Farwell	Mallard	Mallard
Yrs SAGE	6	6	6	6	6	6	8	8	6
Gr. Span	K4-06	E4-06	KG-05	KG-05	K4-05	E3-06	E4-12	K4-05	K3-08
06 Enrollment	487	306	233	242	200	500	442	337	599
% Asian	6.4	.7	2.1	20.2	5.0	8.8	.5	6.5	.8
% Native Amer.	.4	0	.4	1.2	2.7	2.0	1.8	1.5	.2
% Black	80.3	1.3	31.8	31	13.3	3.8	.8	78.6	64.9
% Latino	1.8	1.6	21.9	14	44	2.2	5.9	6.8	19.0
% White	11.1	96.4	43.8	33.5	35	83.2	91	6.5	15
% ELL	1.4	0	13.7	39.9	1.9	8.4	.2	12.5	1.2
% S/Dis	14.1	13	26.7	10.1	9.4	15.4	24	13.8	36.7
% FRPL	84	37.6	68.2	66.1	78.7	64.4	58.4	95.3	88.8
% Gr 3 Reading Prof/Adv	92	96	81	68	74	69	76	57	47
% Gr 3 Math Prof/Adv	76	92	53	52	49	80	76	50	42
SAGE Implementation	15:1 30:2	15:1	15:1 SAGE block	15:1 30:2 Shared Space	15:1	15:1 30:2 SAGE Block	15:1 30:2	15:1	15:1

Table 1
Characteristics of Study Schools 2006-7

At each school, the principal nominated teachers who represented the school's educational practices and we chose 3 teachers as partners for fieldwork across the K-3 span. In year 1, we generated data through eight half-day observations per classroom, standardized environment descriptions, collection of artifacts, and interviews with teachers, principals, students, district administrators, and families. In year 2 we revisited the nine schools to interview participants to gain more understanding of their practice. In year 3 we returned to 3 schools that represented high levels of implementation of the SAGE program, again doing intensive fieldwork that included ongoing observations and interviews. The analysis presented here represents our understandings across multiple years of data collection and engagement in these school sites.

Our analysis followed generally accepted forms of qualitative inquiry, with both inductive and deductive components (Erickson, 1986; Graue & Walsh, 1997). Supported by the qualitative research software NVivo, analysis focused within and across case studies, grade levels, and classroom configurations. Data from the diverse sources were read and re-read, examined through the assumptions that guided the evaluation design and for specific patterns that emerged through fieldwork, with particular attention to assessment. This type of analysis provides the appropriate foundation for transferability (Guba & Lincoln, 1989) from the specifics of local practice to other sites and experiences because it is richly descriptive and comparative. The case studies provide rich information on how, and for whom, class size reduction works and how assessment plays a role in that process. In the next section we provide broad cross-case analysis of assessment in a class size reduction policy implementation and then zoom in with case studies of three schools with high levels of the policy implementation.

Results

Without fail, participants told us that assessment was easier in SAGE classrooms because there were fewer students to assess, or in the case of team-taught classes, a colleague with whom to share the assessment work. Collecting information on 15 students and supervising the rest of the class was seen as more manageable. In addition, teachers felt that the smaller group allowed more effective diagnosis and

intervention—the feedback loop was tight and self-correcting. According to Mrs. Carter, a third grade

teacher at Bethany:

> You're able to see where the needs are and then you can work on those with individual children .
>
> . . . You're able to really individualize so much better. And when you're reading, or anything,
>
> children have so many more turns so you know right away who doesn't know their multiplication
>
> facts, whereas in the larger group you have no clue. So you can't really remediate immediately or
>
> call home or say, "There's this need."

The ability to activate the resource potential of assessments in class size reduction contexts is intertwined

with other demands placed in teachers' instructional practice. With fewer students to work with, many

elements of teaching, including assessment, became more manageable.Meanings of assessment in these

contexts of class size reduction were focused on four elements that uniquely map the current assessment

scene. These four elements, *accountability, alignment, audience, and action*, will be described across all

nine school contexts and then will be illustrated through cases from our most recent fieldwork. They

represent a theme from earlier work—that teaching and learning within classrooms are nested within

systems of practice that include school resources, district policies, goals and tools, and state programs of

all sorts. Rather than being isolated parts of a simple system, the practice of instruction and its outcomes

are inextricably linked to other aspects of education. This is especially so today, with increasing

standardization and shared expectations.

countability

The current education reform context has focused attention on accountability, the idea that

schools, districts, teachers, and students are to be held responsible for the results of schooling. Working

within a system of *accountability*, teachers were challenged to show student growth, particularly in

literacy and math. In all participating schools, staff talked about the force of benchmarks set for student

performance. Even in schools that did not have "testing" until third grade, there was pressure to make

sure students knew the mandated content. Ms. Caster, a first grade teacher at Montford told us:

Everything we do is driven by the frameworks of the testing even though our kids aren't tested

until third grade. We're supposed to be laying the foundation, you know, the basics for them.

And if they don't get in first grade then they are just piling more on in second grade and if they

don't have it then, by the time they get to third grade and they take the test and they don't do

well, then you're judged, your school is judged. So our school is looking at the data but you're

looking at the district and community. They judge the schools by how well you do on the tests,

no matter how much you emphasize that that's a one-day thing. You're still being judged that

way. We're being accountable for what we do.

Although the state accountability program began testing at third grade, accountability was a force in all

schools but in very different ways. The most stringent framework could be seen at Calloway, where the

principal put in place a series of benchmarks toward which staff was to work, biannual teacher evaluation,

and ongoing analysis of all other assessments. Weakest in accountability was West Canton, which had a

general curriculum with a loose systematic framework for assessment or evaluation its efficacy. Teachers

at West Canton were interested in their students' progress and the district had worked toward the states

standards, but accountability was not described as a specter that shaped all their actions.

Accountability was a shape-shifting tool – something that oriented some but that could be ignored

by others. Achievement targets like AYP[2], set to prompt interventions in schools that were falling below

expected levels of achievement, had had official loopholes that deflected attention from less than stellar

performance. This procedural out was ignored at the local level however, with the media highlighting

comparative performance among district schools. Dr. Post, the principal of small semi-urban McMahon

Elementary, talked about the perceptions and practice of accountability testing in his district:

Everybody did pretty bad in math. AYP -- 47.5% is the cut off. We were at 47. So both our

reading and math we were point five under, but we're exempt from being put on the *Schools In*

---

[2] Adequate Yearly Progress (AYP) is defined by *No Child Left Behind* as the improvement of student performance relative to the lowest achieving school or demographic group over a set time period. Each year schools are to attain higher levels of achievement until at the end of 12 years, all students are proficient on state assessments.

Need of Improvement list. Because the number of our kids is so small, under fifty, you, they

don't count you. . . .Race, white kids did, I think 79%, black kids were like ooh, it was way down

there, 25%, so it's a big gap on math based on race. And SES was a big one too. The sample size

is so small it's only based on 36 kids, 47 kids took the test, only 36 were full academic year, but

to come off that 81% the previous year. Unfortunately that's how you get judged and that's what

hits the papers, and then you try and explain it to people and it makes it sound like you are

making excuses.

Accountability focused attention on student performance, forcing school people to examine outcomes of

their efforts. This focusing could be productive, prompting assessment of alignment and reorganization

of effort to improve curriculum through teacher capacity, or it could be overly focused on test results

without attention to productive teaching. Balancing attention to both inputs like teacher knowledge and

skill and output like test scores seemed to be the most productive approach. The process used to build

that accountability system is detailed next.

Alignment

Schools worked towards accountability through various forms of *alignment,* the degree to which

curriculum, standards, assessment, reporting procedures and instruction form a coherent system. With

increasing pressure to document student progress, teachers juggled intense demands. Assessment often

served as a bridge among the elements and those who worked in an aligned system found their jobs less

fragmented. For example, at Montford, Mrs. Durst, the principal, described how they worked through

alignment:

Last year we found out was that kindergarten wasn't doing math Trailblazers. First grade, second

grade were. But throughout the district there were some inconsistencies. Kind of like a

smorgasbord. You could choose what you wanted to teach when you wanted to teach it. And

especially with the impact of the standardized testing and all the frameworks we got from DPI,

we needed some consistency. So Mrs. Felton and Mrs. Monroe both worked at the district level in

creating a guide for K through 2 in monthly goals of what they should be teaching. In math. So

these are the have-to's, these are the can-do's but they have to do the have-to's. And it will lead

towards, hopefully this year the same group's going to meet together creating some common

assessments which is another step the district will be taking.

Developing shared expectations for teachers and students made alignment a powerful community building

experience. Teacher participation in the alignment process was key to professional buy-in and made it

more likely that instruction connected to assessments. Things were more difficult in schools that lacked

alignment – the system was more chaotic, the assessments seen as more of a burden. At McMahon, the

teachers resisted an increasing number of district-mandated assessments and the lack of coherence among

the curriculum, tests, and report cards. This was especially difficult for the most senior teachers who felt

that the new assessments represented a lack of trust in their professional expertise. Nothing seemed to fit

together and though having fewer students made the job more manageable, they now felt at sea. Mrs.

Ludwig at McMahon was frustrated because her preferred reading and assessment method, Accelerated

Reader (AR), was not valued by district administration, who had mandated the use of an assessment

named On the Mark, a comprehensive package of formative assessment. She saw the district move as one

that privileged consistency across the district over individual teacher judgment:

> You read books where you think they are, and you assess by different forms of questions, which I
>
> do in AR. And you can't convince administration that with AR you can do your guided reading
>
> groups, you can do anything you want with AR. You can work with them, you can say, "Does it
>
> sound right, does it look right?" You can do all of the politically correct things with AR but they
>
> only recognize *On the Mark. On the Mark* takes such a long time.
>
> Graue: And they want you to do that so they have?
>
> Ludwig: Consistency.

These two quotes illustrate the tension in alignment – when alignment was a useful tool, it provided an

effective structure in systems that had a shared vision for practice and learning. When it was not helpful,

it was seen as an imposition of power that lessened teacher authority. The key seemed to be hitting that

sweet spot between the general and specific, between shared goals and individual intentions.

Audience

At the heart of the alignment problem was the issue of *audience*. In the assessment driven systems of today, assessments serve varied audiences for different purposes. Sometimes assessments are designed to inform classroom decision-making, sometimes they are used to track school efficacy. The farther the assessment's audience from the classroom, the less likely it was to be seen as helpful to instructional practice. A good example of this is the use of the DIBELS in a large urban district to track student acquisition of phonological awareness, alphabetic understanding, automaticity and fluency. A benchmark tool, DIBELS scores were used to track the efficacy of curriculum and to describe patterns in achievement in large subgroups of students. Teachers saw the measure directed to an audience outside the classroom because it seemed irrelevant to their instructional work. It was necessary then to supplement DIBELS with other types of assessment for various purposes. Mrs. Feller, a first grade teacher at Bethany, noted:

> For formal assessment for reading is the DIBELS assessment in first grade. Although what we
> didn't like about that was it didn't give us anything real concrete to show our parents about
> reading because it would just be nonsense words. So I actually go and I test the kids 3 or 4 times
> a year on like a text level analysis, word list, because the parents really wanted something more
> concrete.

When assessments were required for outside audiences and teachers could not see the relationship to their own instruction, their practice felt unaligned. Teachers at Gallows described a bewildering number of assessments needed to satisfy district and federal audiences for the schools' grant-sponsored programs. While some were useful in their instruction, many were solely to inform the evaluation and teachers balked at using valuable classroom time for assessment that did not inform their teaching. For assessment to be seen as a *part* of instructional practice, something that was not an added burden in a very busy schedule, teachers needed to see the practical connected to their work. This was the case at Earhart, where the district had developed assessments in literacy and mathematics that provided both district level evaluation information and classroom level instructional information.

Action

The final element was *action*, the degree to which teachers felt they could *do* something with assessment information. Mrs. Walworth, the principal at Earhart, instituted a practice where teachers regularly met with reading specialists studying student growth through assessment walls. All student assessment information (running records, text levels, etc.) was posted on a wall and teams reviewed student data for trends over time and necessary instructional interventions. Teachers helped each other problem-solve and students moved among instructional groups so teaching matched current skills and needs. The assessments took on heightened importance because they were integral to the instructional planning.

Action was not always positive. The pressures teachers felt to have children perform on assessments sometimes pushed them to resort to bribery. Mrs. Manchester, a kindergarten teacher at Wellstone Blvd., needed children to count to 100 by the end of the school year. An initial assessment pass found 3 counters in her class of fifteen. She made it more likely by making the following announcement to her students:
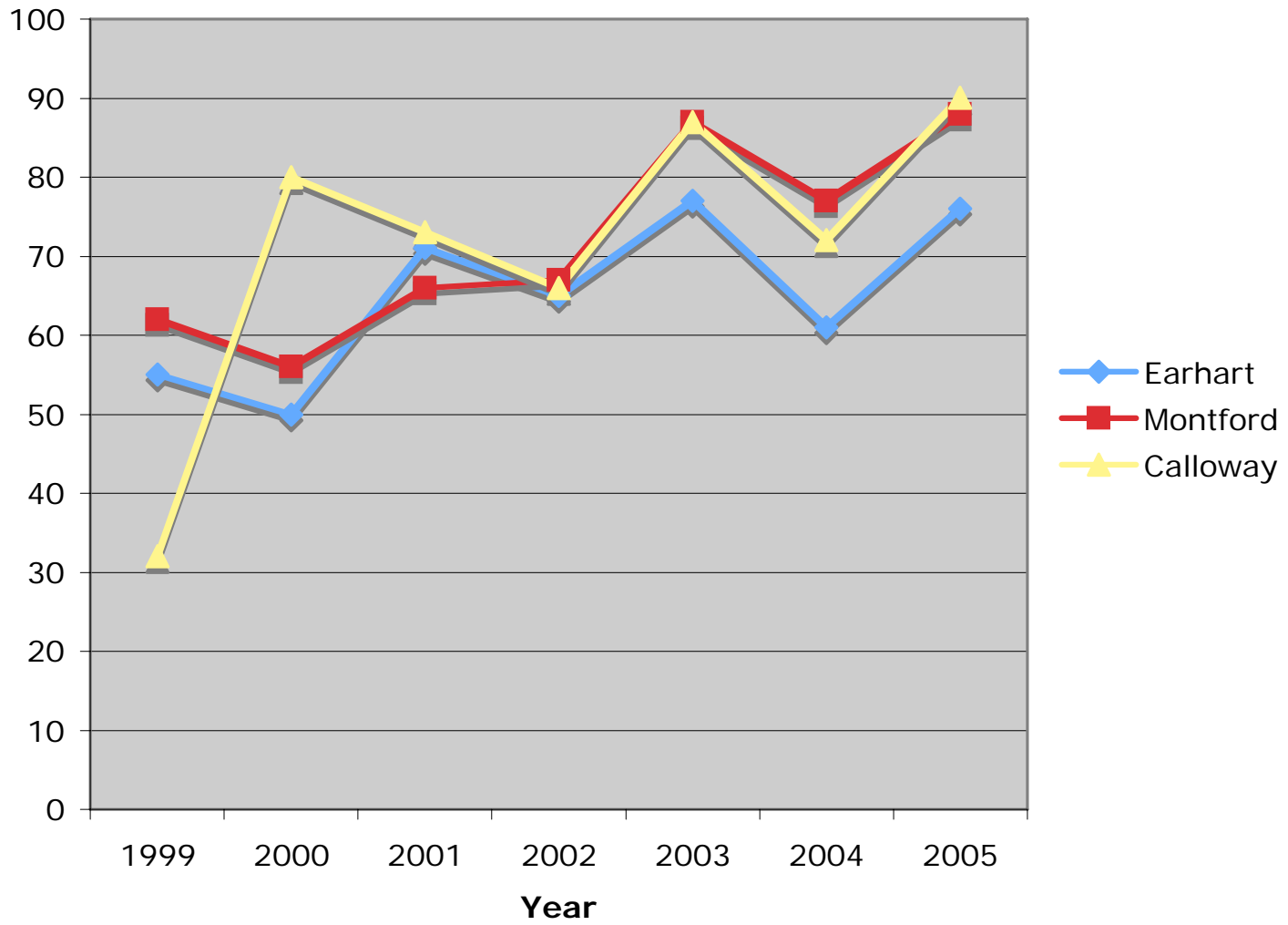
> Last week I tested people who were ready to count to 100, and Miguel, Thuyet and Jared counted
>
> to 100. And those three boys are going to get a candy bar at the end of the day. (Kids make a
>
> slight gasping sound and look around group). I went out and brought candy bars for the people
>
> who counted all the way to 100. And tonight you're all going to get a sheet to practice your
>
> counting, and when you're ready to take the test, if you count all the way to 100, I'm going to
>
> give you a candy bar!Focused attention on an assessed need can jumpstart student learning,

particularly when it is followed by targeted instruction. In this case the action of a sweet incentive and outsourcing the learning to the home might have missed some in-class learning opportunities that could enrich the education in this class.

Accountability, alignment, audience, and action were themes that we identified in the nine-school assessment context across three years of data collection. They represent a perspective that situates teacher action and student outcomes in nested contexts, related to each other and to other elements of the

schooling context. Their importance to the implementation of SAGE rests on the recognition that SAGE

teachers work within systems that are more complex than just class size reduction—their instruction is

embedded in teaching systems that include complex attention to assessment.

To this point, our discussion of assessment and class size reduction has not been explicitly

synthetic; examining the power that each has to activate the resources possible in the other. In the next

section we illustrate their connections through three case studies that demonstrate the synergy in

assessment and class size reduction. The three schools represent a subset of our larger study that we

identified as having high levels of implementation of the SAGE program, with strong practices in each of

SAGE's 4 pillars of class size reduction, rigorous curriculum, professional development, and home school

relations. We provided an overview of the schools in the context of the broader study in Table 1. Figure

1 illustrates their achievement over time:

WRCT 1999-2005 subsample

Legend:
- Earhart
- Montford
- Calloway

X-axis: Year (1999, 2000, 2001, 2002, 2003, 2004, 2005)

Y-axis: 0 to 100

Negotiating Vision: Building an Assessment Framework at Calloway Academy

Calloway Academy is a mid-sized school situated behind strip malls and chain stores on the outskirts of a large urban area. The past ten years have brought steadily declining enrollment and shifting demographics. While the percentage of white children has remained relatively consistent, accounting for half of the school's population, the percentage of children classified as Hispanic has doubled and the percentage of African-American children has dropped by fifty percent. Though initially minute, the percentage of Asian students has also increased more than four-fold. In 2005-06, approximately 70% of Calloway's students were classified as economically disadvantaged. Traditionally serving students in pre-K through fifth, in 2006 Calloway began a transition to a K-8 structure. This move is at once an effort to delay, and hopefully circumvent, the need to merge with another area school, and a response to parents' desires to keep their children in a smaller, more nurturing context instead of sending them to middle school.

Calloway entered the SAGE program in 2000. All K-3 classrooms have 15 students and 1 teacher, with the exception of one 30:2 second grade in which the teachers chose to partner. With the pending addition of seventh and eighth grades in the next two years, the current SAGE configuration is likely to shift as classroom space becomes scarcer. Three veteran teachers participated in our research. Linda Trainer teaches 15 kindergartners in a spacious classroom with large windows and a high ceiling. Gloria Howard and Marsha Delton teach first and third grades, respectively, in adjacent classrooms. Though Marsha had a larger room last year, she recognized the impending space crunch inherent in adding grades 6-8 and chose to make the move before she was forced to do so.

Calloway could be described as a traditional school. Speech therapists, special education teachers, psychologists and those who teach the newly implemented programs designed to aid English Language Learners, collect children at predetermined times and take them to other rooms to work. All of the teachers involved in the research project have some form of seatwork and reading group time in the mornings, with music, physical education, library/computer lab time and art sprinkled throughout the

week. Students have two short recess periods, one in the morning and one following lunch. A newly

implemented breakfast program has children eating in the classroom for approximately fifteen to twenty

minutes in the morning. Teachers piece together planning time during these activities. Recent union

negotiations have led to an optional twelve hours per year of collaboration and planning time, for which

teachers must be compensated.

Assessment practices must be interpreted in the context of ongoing labor issues that dictate the

amount of time teachers engage in professional duties. Because, contractually, teachers must be

compensated for collaborative planning time and the principal may schedule, at maximum, two hours of

faculty meetings per month, teacher time is at a premium. While some teachers are willing to stay for

extended periods to plan events, work with colleagues, and participate on committees, others see these

activities as being beyond their contractual obligations and choose not to participate. However, perhaps in

light of these challenges, the school's principal, Mrs. Collier takes much of the responsibility for

alignment and accountability on herself, through coordination and oversight of teacher collaboration,

school, classroom, and student level curriculum planning, and on-going data disaggregation.

In response to poor performance on the math component of the annual state examination, Mrs.

Collier spearheaded a plan for math pre and posttests at each grade level and an extended two-hour daily

math period for all students. Additionally, she required teachers to provide extra help to one or two

students who need it during one of their semiweekly music class periods. She holds biannual monitoring

conferences with each teacher, asking detailed questions about the student progress, strategies for meeting

the needs of struggling students, discipline concerns, and the additional support they feel they need.

Through these conferences, Mrs. Collier presses teachers to analyze their students' needs and the

instructional strategies used in response. An example of the documentation for these conferences is

provided below:

Teacher: _____
From: (principal)
Re: Monitoring Conference

A monitoring conference gives us an opportunity to discuss the progress students are making, the progress made toward covering the Learning Targets, and to share ideas.

The following should be brought to the conference:
1. Assessments from the reading, language arts, social studies and science
2. Evidence of your alignment of what is taught in your classroom to Learning Targets
3. Evidence of students' rate of success in the various areas.
4. Anything else you want to share
5. Make a copy of your responses for me, and you keep the original.

Please jot down a few notes to these questions prior to our meeting. <u>Make a copy for me</u> and keep the original.

Reading

1. What percentage of your students is below level? (teacher crossed out "below level" and wrote "*still learning the alphabet*")
Lisa, Billy, Jenny, Cory
What percentage is above level? (teacher crossed out "*knows all K wrds or at readiness B or L*"
7 students

2. Who are the struggling students? How far below level are they?
Hugo Gleason

3. What is being done to help the struggling students? What strategies are being used? What additional help is being given? How often is additional help given?
-- guided reading groups, small group instruction—at least 3x per week
-- students w/ with Mr. Apple
-- vocabulary w/s—sent home w/homework packs—also letter cards
-- Alpha friends in reading series teach the letter sounds

4. What are you doing for phonics?
Alpha friends, phonics library stories, phonics center kit

5. Have you made any changes to the kinds of questions you ask?
Explanation for the answers/ discussion of answers/ illustrate answers

Teachers are required to submit six writing samples for each student throughout the year, which Mrs. Collier evaluates. Mrs. Collier, in collaboration with the school's learning team and textbook adoption committees works to ensure that newly adopted textbooks align with the district's learning targets (or benchmarks). At the district level, this large, decentralized administration is working toward alignment,

by designing cross-district, publicized learning targets and fostering formal partnerships with local

universities to design classroom assessments that align with these benchmarks as well as guides for

"analyzing and learning from student work."

Calloway has all of the pieces in place for instruction-driven, rigorous assessment practice that

leverages the benefits of class size reduction, but something is missing, namely, inclusive and proactive

coordination across and within classrooms. Most teachers have not embraced the need for concentrated

effort in this area, feeling most accountable for the children in their own classrooms and blaming

shortcomings on students' home lives. Though most assessments, with the exception of DIBELS, a

district required phonemic awareness test, are seen as tools to inform the action of the classroom teacher,

record keeping and coordination across classrooms is only beginning to take form. For example, as a

result of the first math pre-test in kindergarten, the teachers learned that the pre-K teachers were not

teaching some math concepts others assumed were being taught. This resulted in an informal "hallway"

conversation, and the pre-K teacher began introducing number words so that they were more likely to

have the knowledge for the kindergarten math assessment.

Although Mrs. Collier seems to have assessment, accountability, and alignment practices in mind

when she holds monitoring conferences, some teachers view these as an insult to their professionalism.

What is needed is buy-in beyond the parameters of the teachers' contract, facilitated in the fickle area

between the leader and the led. While the principal's ideas may be on the right track, many teachers felt

voiceless at Calloway, unable to shape decision-making or to make their needs known. This is how Mrs.

Delton described it:

> I like the small school but I think it's not being run in a way that gives people a voice. Because I
>
> thought if I'm in a smaller school (because I came from a huge school, lots of staff, lots of kids),
>
> and when I came to a small school I thought I'll have a voice. And I think only a few people have
>
> a voice. Even on the learning team, I didn't have a voice. I don't care what anybody says, I didn't
>
> have a voice. So I'm kind of disenchanted with the building in general.

Disenchantment on the part of a teacher is a disadvantage when working to create a systemic and effective accountability system. They are on their way, but will need investments of time and effort to make the next grand leap forward.

Assessment as a Responsive Practice – Montford

Montford is a rural school located in a river town of 18,500. The school enrolls 500 children in grades K-6. The majority of the children are white (83%) but there are also Hmong (9%), Latino (2%), African American (4%) and Native American (2%) students making this school a diverse one relative to the town where 93% of the residents are white. Eight percent of the children at Montford are English Language Learners. The school's free and reduced lunch rate has been growing over the last five years and is now at 64%. Montford implements SAGE through 15:1, 30:2 and SAGE block[3] configurations.

The school adopted the Responsive Classroom Model (Northwest Foundation for Children, Inc., 2007) six years ago and it is incorporated into all facets of teaching and learning. The Responsive Classroom values both social and academic learning and recognizes that process matters as much as content. From this perspective, assessment practices are seen as social interactions that take into account the needs and experiences of all stakeholders.

Mary Durst has been Montford's principal for 5 years. She believes she is a "fine-tuner" of assessment practices and she makes sure "it's not just something that fills up paper, that we're actually using that information to change what we are doing with kids." At a recent shared leadership meeting, a representative from the kindergarten team worried that the benchmark testing at the beginning of the school year would be better timed in November when students were more familiar with school routines. Mrs. Durst referred this issue to the school's action team who decided that the *when* of testing is sometimes just as important as the *why* of testing. Testing was shifted to November because it made sense to address the children's needs.

.[3] SAGE block configurations provides an additional teacher for the reading and math periods

Dena Monroe co-teaches a SAGE second grade half day and is the school's professional development coordinator the other half. In this role she helps teachers find ways to meet their students' needs. The link between this goal and assessment is what Mrs. Monroe calls "informed practice" and what kindergarten teacher Kathy Martin suggests provides an opportunity: "we can see where they are, where we need to work with them. It helps us to make sure we are teaching what we're supposed to be." At Montford these are calls for responsive action.

Karen Martin and Nancy Giles are both veteran kindergarten teachers who engage in assessment as instruction and instruction as assessment. Their practice includes daily news where children "fill in the blanks" about the news of the day. Both teachers observe how the children are doing and work with individuals during learning stations to address needs they noted during the daily news session. Mrs. Giles might suggest that Maureen "write the room" and copy down all the words she sees. Mrs. Martin might work with Chue and James on alphabet recognition.

Mrs. Felton is a first grade teacher who has decided that she could do better writing instruction. She has embraced Lucy Calkins' Writing Program (Calkins, 1994) focused on writing as a joyful and creative process. Everyone is learning the pleasures of writer's workshop and the "writing the small moment." During a recent lesson, Mrs. Felton sat next to Walter, who is slowly transitioning from a self-contained special education placement to Mrs. Felton's class. Mrs. Felton and Walter talked about how he could draw pictures of the words he could say but had yet to learn to write. Mrs. Felton used the information from these one-on-one sessions to shape the next day's lesson for the group. Mrs. Felton also assesses her practice as a member of a team of teachers who are new to this program and understand that change isn't always easy. These teachers meet regularly to discuss the challenges and triumphs of this new curriculum.

This team approach to problem solving was a common one at Montford. Mrs. Monroe and her partner, Mrs. Bonkowski, decided that some children are struggling with double-digit subtraction. The following day Mrs. Bonkowski continued with the lesson as Mrs. Monroe worked with small groups reviewing the previous day's problems with each child. Mrs. Monroe was able to provide "just in time"

instruction that addressed needs on the spot rather than deferring them to later action. And this "just in time" idea extended to scheduling which includes the flexible use of an instructional aide who worked one on one with a student in the class. At times Mrs. Hester worked with Sam in the classroom but when the content included concepts that he struggled with they worked outside of the classroom. Scheduling to maximize resources at a specific moment in time is characteristic of the school.

The staff at Montford is committed to the Responsive Classroom's call towards "knowing the families of the children we teach and working with them as partners is essential to children's education." (Northwest Foundation for Children, Inc., 2007). Recognizing the importance of a shared vision for learning, Montford holds *hopes and dreams conferences* at the beginning of each year where families are invited to share their goals for their children. Usually these "hopes and dreams" are displayed in the classroom as a reminder of where the journey began and a shared vision of their destination.

Alignment is what Mrs. Monroe calls the "silver bullet" – an assessment package that would provide real information for all the stakeholders: the state and district, teachers, children and families. Mrs. Monroe and a team of teachers have just completed a draft of the *SAGE/District Assessments Matrix*, which synthesizes the Wisconsin Frameworks for Reading and Math (which inform the state assessments), SAGE goals and the school's own reading and math benchmarks. This aligns the district requirements using WKCE scores as the standard of success, SAGE requirements, and classroom practice guided by an expert teacher. At the school level, kindergarten teachers began the process of creating benchmarks, passing them on to their first grade colleagues, with each grade level having an iteration. It is this sense of building from the foundation that is central in Montford's quest for improvement.

Assessment practices at Montford are a work in progress. Mrs. Monroe knows that the report card isn't aligned with the curriculum and needs to be changed. The new reading curriculum includes tests that the teachers feel are of limited value relative to their time requirements. The math curriculum doesn't allow for re-teaching once a unit test has been taken. It is designed to move on after each chapter test with the understanding that concept reteaching will take place at a later time. This seems to run counter to the philosophy of these teachers who have a keen eye for teachable moments and instruction that comes

sooner rather than later. However, there is much forward movement and it is shaped by thoughtful

consideration of the learning needs of all children, the goals of families, and the professional work of

teachers.

Establishing common expectations through alignment & collaboration

Earhart Elementary

Earhart Elementary is a small diverse K-5 school nestled in the middle of a working class

neighborhood in a semi-urban school district. More than two thirds of the students are classified as poor,

1/3 are English Language Learners (split between Latino and Hmong), 1/3 African American, and 1/3

white. The school has participated in the SAGE program since 2000, and now has K-3 implementation at

15:1 (with one pair of classrooms a 30:2 shared space context). Three teachers participated in our study

in 2006-7. Molly Masters teaches kindergarten in a small but exquisitely designed space shared with

another kindergarten in a 1000 square foot classroom.[4] With a "wall" of cabinets separating the two

classrooms, the teachers synchronize their schedules so that the joyful noise of kindergarten learning is

simultaneous. Tammy Helman works in a spacious classroom with 12 first graders. Lauren Rich teaches

15 second & third graders and does some teaming with her partner Betty Miller, whose classroom is next

door.

Schedules at Earhart are designed to maximize instructional support within the classroom.

Instructional resource staff such as Title I reading and ESL teachers, have planning time with classroom

teachers and work within classrooms to link their instruction to the core classroom curriculum. This

results in sliding student/teacher ratios that range from 15:1 to 12:3. Instruction in each of these

classrooms is mixture of large group, small group and individual teaching with teachers blending

instruction, assessment, and evaluation in a tight loop. This loop is facilitated by the district's progressive

alignment of standards, curriculum, assessment, instruction, professional development, and reporting

formats. Beginning with the content standards, the district has worked to develop assessments that will

---

[4] This is shift from the 2004-5 school year when 2 teachers team-taught in this space with more than 30 kindergartners. The difficulty of managing that many five year olds motivated the shared space solution.

inform both instruction and district decision making prior to state testing in third grade. The teachers find

the K-2 Literacy Instrument particularly well suited to the kind of daily instructional decisions they need

to make. The K-2 Math Tasks are less informative for formative assessment because they are given only

twice a year and students are to be tested to the grade level standards at the beginning of the year, which

teachers find highly frustrating for many students. Ms. Helman saw the Math Tasks as what Shepard

would call benchmark assessments, less related to instruction than programmatic evaluation:

> I don't think the district would say this. It's part of measuring our progress and the quality of our
>
> teaching and how our students are achieving and learning and whether we're closing the
>
> achievement gap. It's probably to inform the public how the school system is doing.

These assessments are linked to instruction that is supported through ongoing professional

development embedded within the school day.[5] The district has gone so far as to develop grading guides

that link assessments, standards and report card proficiency levels so that communication of student

progress is aligned seamlessly with assessments. Ms. Masters, a second year kindergarten teacher, works

to coordinate her assessment knowledge with her practice by going over student work daily and

documenting their progress in content based assessment binders. She sets up centers that mirror the

requirements of the report card that reflect the grade level standards. She is the paragon of alignment,

working to connect all aspects of her practice. Even in a context of close alignment, the official

assessment results can have a jarring effect on teachers' psyches. Ms. Helman told us that she needed to

prepare herself emotionally for student performance even though she did close assessment with her

students on a daily basis. The tasks on the K-3 Literacy Instrument contrasted strongly from the books

she uses in instruction "so the results can be *horrifyingly* different (laughs) sometimes. . . Sometimes this

is a real wakeup call. I guess I feel like they need to be able to be successful on the assessments like the

K-3 Literacy Instrument and the K-2 Math Tasks." When asked what she does when the results of the

---

[5] This embedded staff development approach, favored in the literature for improving practice, is becoming increasingly rare in the district with severe budget cuts curtailing any support viewed as discretionary.

assessment diverge from what she thought a child could do she simply said, change her practice. Rather than relying on her assessment of learning levels, which is developed in scaffolded instructional settings, Ms. Helman needed to look at what students can do without teacher support.

> When Destiny couldn't pass the 7, I thought, OK that doesn't feel like a reflection of the child I normally see. But this is what she did, so I'm going to have to take action based on it. I'm going to have to adjust my teaching. I'm probably going to do with her what I did with Matthew. For the last three weeks or so I met with him a second time and I gave him extra books and just really worked on using the reading strategies. I feel like I can see that Destiny isn't doing that independently without my prompting and she has to be able to do it by herself, so I need to get her there.

Communication of the targets of instruction among staff but also between teachers and students provides a strong community sense of the goals, grounding instruction in a sensemaking that has purpose. Student self and peer assessment was regularly used at Earhart as a way to communicate the goals teachers were addressing. In Ms. Masters' kindergarten, students have *I am working on* sheets in their writing binders that list skills that students are to practice in their writing, like using 2 finger spaces. They are charged with assessing their success in achieving the goal, and asked to provide evidence of their accomplishments with writing conferences with Ms. Masters. In Ms. Helman's class students provided feedback on writing, with their teacher modeling the connections between the strategies they were learning and student use. The richness of the conversations around peer evaluation were made possible, in part, by the small group size of twelve first graders and Ms Helman's modeling of supportive evaluation:

> After Kiera read her book about music class, she received the following comments:
>
> Alison: I like how you did your pictures and how you were about music and I noticed you went
>
> > "boom and boom and boom."
>
> Ms H: What craft is that called? Using sound words!
>
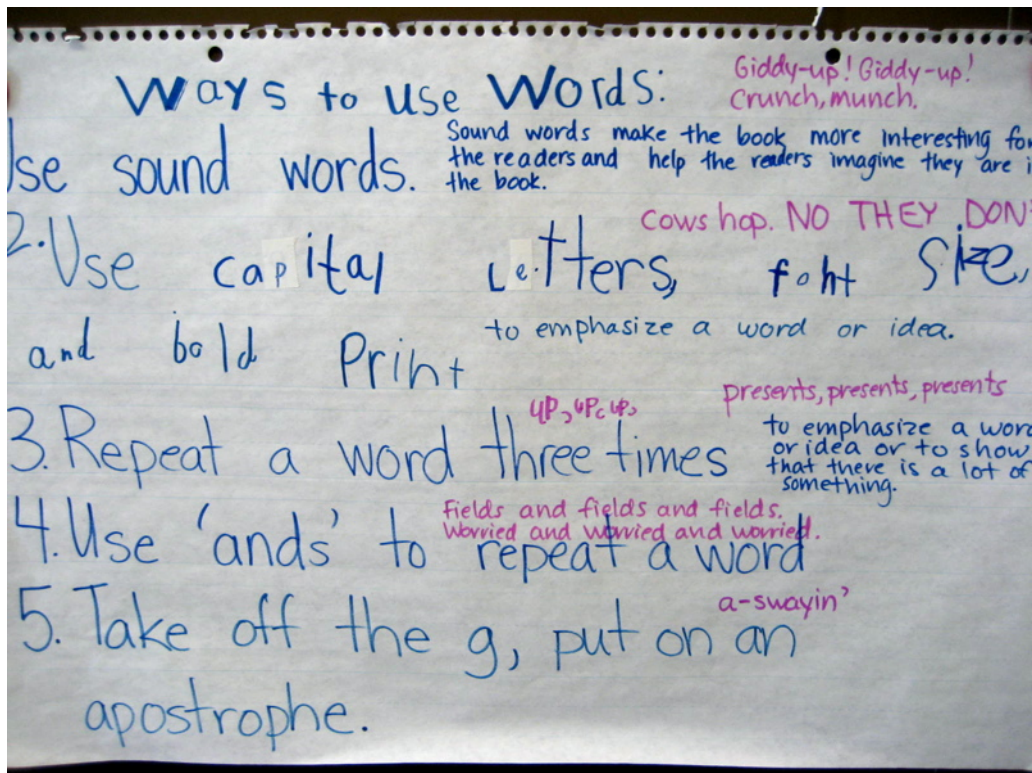> Wendy: She said, "I like music very very much."

Antoine: I like how you took your time and you didn't scrabble.

Ms H: What craft did she use that you can see with your eyes? You can see it by looking. Hold

up your book and show them"

Kiera holds up her book and someone replies, "Bold print."

Ms H: Were you emphasizing your words? So you used bold print.

This public sharing of criteria for performance and evaluation was done in a positive collegial manner and

it referenced instructional materials generated by the class that described writing crafts, seamlessly

connecting instruction and assessment:

This self assessment is a practice-based example of Black and Wiliam's assertion that:

Self-assessment by pupils, far from being a luxury, is in fact an essential component of formative

assessment. When anyone is trying to learn, feedback about the effort has three elements:

recognition of the desired goal, evidence about the present position, and some understanding of a

way to close the gap between the two. All three must be understood to some degree by anyone

before he or she can take action to improve learning. (Black & Wiliam, 1998, p.)

In her third year as principal at Earhart, Paula Walworth has been leading her staff through a comprehensive school reform, Professional Learning Communities (DuFour & Eaker, 1998), focused on shared leadership and professional collaboration. PLC is centered on three essential questions – 1) What do we want students to learn, 2) How will we know if they learn it, and 3) What will we do if they don't? These questions and their action-oriented premise shape much of the work at Earhart. The staff has worked to develop shared expectations for both academic and behavioral learning and to communicate them clearly to students, staff and families. These shared expectations are communicated in common language and rules, common standards for learning and behavior and multilingual communication. Grade level teams have scheduled weekly collaborative time and these groups meet periodically with the instructional resource staff that supports them. Each grade level team has a representative to the shared leadership team, the building decision making body, which meets monthly. A commitment to professional learning community does not mean that there is not a role for leadership from the principal. Mrs. Walworth has a vision for her staff, a vision that is based on collaboration and shared commitments to instructional improvement which included developing common assessments at the school:

> So there're people that have self-selected, they want to be part of this process. But the question will come out, "Well what's going to happen you know for people who aren't working with us to create common assessments? Who aren't willing to change their instruction because some kids aren't successful?" And my very honest answer is "then they may find this is not the right place for them to stay."

Both assessment and class size reduction are implemented at Earhart through collaborative, systemic practice. Rather than assuming that individual teachers can/should generate learning or reform, staff work together through consultative relationships to maximize the resources for teaching and learning. The district's work to align standards, curriculum, assessment, and reporting has provided a system that gives teacher work a more coherent quality than at other schools and relies of teacher professional knowledge that is scaffolded by appropriate tools. Strong knowledge of content has been generated through staff development and connected assessment practices that inform instruction. As is

the case at Montford, the weak link could be seen as the report card, which appears to serve a benchmark evaluation purpose rather than a tool to communicate with families. The impenetrable, standards-based language is difficult for even the most educated parent to understand and is exacerbated by a push at the district level for staff to limit open-ended comments that are seen as indicators that parent-teacher conferences are needed. In working to communicate to a district audience, the report card is losing its community focus, an indication of the fragility of aligned systems.

Discussion

Resources like class size reduction and assessment can be catalysts for positive change if they are activated in thoughtful ways. These resources, used in ways that increase understandings about children's learning and that empower teachers and principals to use that information relevant in local contexts, create opportunities for teacher support for learning.

Previous work on class size has indicated that in smaller classes, students have more access to teacher time, are taught in smaller groups, and have their learning disrupted less frequently by disciplinary issues. Smaller classes provide opportunities for teachers to spend more time per student for all teaching's complex interactions– instruction, assessment, management, and social and emotional connection. These opportunities are realized in contexts where structural commitments facilitate teacher action that is responsive to the needs of their students. Classroom spaces must be sufficient, teacher workloads reasonable, intellectual and material resources available to the task at hand, and shared commitments among colleagues perceptible.

Our work with SAGE teachers demonstrates the promise of linking smaller classes and assessment through accountability, alignment, audience and action. Although accountability is often depicted as a conservative force designed to dictate teacher action, it can provide an intentionality that makes teaching more effective. We are reminded of Lilian Katz' principle of optimum effects – good for children is good only in the appropriate proportions. So it goes with accountability and its related elements. In systems with loose models of accountability there is little direction for teachers and learners and "if you don't know where you are going, you'll probably end up somewhere else" (Berra, 2002).

Conversely in systems with overly strict, top down accountability, there is little place for teacher thinking

and problem solving – they lose their voice as the teachers did at Calloway. In its most extreme form,

accountability is fused with high-stakes testing.

Examining assessment practices force us to look at the needs of three distinct levels of education:

the district, the school, and the teacher. This is necessary because assessment is undertaken by individual

teachers in response to school programs that are nested within district policy and practice. Our challenge

in this project has been that SAGE is a state focused program, implemented in individual schools, but

lived in particular classrooms in very specific ways. The relationship between student achievement and

assessment practice is not a linear one, which is not only because of a methodological mismatch within a

study. It uniquely reflects the connected nature of assessment in today's schools, which come out of

federal and state mandates, curriculum needs, and professional knowledge of teachers and administrators.

Assessment practice would seem to be facilitated by smaller classes; with fewer students there are

more opportunities to assess and to teach in response to assessment results. High quality assessment

practices were more likely to take place in schools with higher achievement and there were explicit

connections *through* assessments to the varied communities of interest: the district, the school, the

teachers, students, and families. These connections increased the likelihood that the assessments met the

needs of these audiences and that they prompted some kind of action. Action was the linchpin of high

quality assessment. In good assessment systems, teachers had tools that they understood and that they

could use to improve practice. This improvement was a contingent one, related to the needs of their

students this year. In contrast, assessment in lower achieving schools took place in disjointed systems

that focused primarily on summative rather than formative assessment. In these schools, teachers had

tools to find out where students were but this knowledge was not connected to instructional action.

High quality formative assessment, which informs instructional practice in the moment to

moment of teaching requires teachers to work a delicate balance, one that takes in information about their

students while guiding their learning. Rather than being separate, instruction and assessment have a

seamless, recursive quality that makes it difficult for outsiders to discern when one strategy shifts to

another. It broadens assessment beyond the classroom teacher to include students and colleagues, recognizing assessment as a communication strategy that can include many in the conversation.

High quality teaching in SAGE schools was more than a collection of strategies of best practice. Although strategies are important, they are not enough. Effective teaching in SAGE schools had a systemic quality that linked the use of resources, relationships among staff, collective goals and tools to accomplish them. It recognized that teaching and learning reach beyond the classroom because the forces of education do as well. For class size reduction and assessment to be effective resources, we need to move beyond classroom-based solutions to approaches that align the goals, practice, and tools of education in ways that support teacher professionalism and clarify intentions for all participants. But all this alignment must not come at the cost of being responsive to the needs of our youngest stakeholders, namely, the students. A good educational plan should not crowd out attention to the needs, experiences and interests of students. A dialogic relationship exists between where we are and where we want to go in the journey of teaching and learning. Smaller class sizes and formative assessment makes it more likely that teachers can orient themselves in the accountability map in relation to their students.

This work is itself an assessment activity, prompting us to recognize where we are in our understanding of SAGE and where we want to go in the future. From it we make the following suggestions for future research:

- To most adequately address the practice-achievement link, we need to have classroom level data for sampling and analysis. Without it, our sampling focuses on the school level, based on aggregate achievement while our data collection is at the classroom level.

- We need to examine carefully the answers on question 9 of the teacher survey that addresses assessment opportunities and relate those answers to our case study knowledge in the nine-school sample and the instructional quality ratings on the CLASS to see how they provided convergent or divergent portraits of assessment practice. If they diverge we need to rethink the teacher survey items. This comparative analysis can be extended

to the implementation typology analysis that will use these elements along with student

demographic characteristics and historical achievement information.

- We also need to examine the potential of the End of Year Report to provide information

    on the alignment of standards, curriculum, instruction, assessments, and reporting

    formats. Given the importance of alignment in this work, it would be helpful to see how

    the current tool provides information and to consider how we might include additional

    questions to support our knowledge.

References

Biddle, B. J., & Berliner, D. C. (2002). Small class size and its effects. *Educational Leadership, 59*(5), 12-23.

Black, P. & Wiliam, D. (1998). Inside the black box: Standards through classroom assessment. *Phi Delta Kappan*

Blatchford, P. (2003). *The class size debate: Is small better*. Maidenhead: Open University Press.

Blatchford, P., Moriarty, V., Edmonds, S., & Martin, C. (2002). Relationships between class size and teaching: A multimethod analysis of English infant schools. *American Educational Research Journal, 39*(1), 101-132.

Cahen, L. S., Filby, N., McCutcheon, G., & Kyle, D. W. (1983). *Class size and instruction: A field study*. New York: Longman.

Calkins, L. (1994). *The art of teaching writing*. New York: Allyn & Bacon.

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis, 25*(2), 119-142.

Darling-Hammond, L. (2004). Standards, Accountability, and School Reform *Teachers College Record* Volume 106 Number 6, 2004, p. 1047-1085, http://www.tcrecord.org ID Number: 11566, Date Accessed: 3/15/2007 6:41:06 AM

DuFour, R. & Eaker, R. (19998) *Professional learning communities at work: Best practices for enhancing student achievement*. Bloomington, IN: Solution Tree.

Ehrenberg, R. G., Brewer, D. J., Gamoran, A., & Willms, J. D. (2001). Class size and student achievement. *Psychological Science and the Public Interest, 2*(1), 1-30.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119-159). New York: Macmillan.

Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide

experiment. *American Educational Research Journal, 27*(3), 557-577.

Finn, J. D., Gerber, S. B., Achilles, C. M., & Boyd-Zaharias, J. (2001a). The enduring effects of

small classes. *Teachers College Record, 103*(2), 145-183.

Filer, A. (2000). Socio-historical & cultural contexts of assessment policy In A. Filer (Ed.),

Assessment: Social practice & social product (pp. 7-10). London Routledge Falmer.

Glass, G., & Smith, M. L. (1979). Meta-analysis of research on class size and achievement.

Educational Evaluation and Policy Analysis, 1(1), 2-16.

Graue, M.E., Hatch, K. & Rao, K., Oen, D. (in press). The wisdom of class size reduction.

American Educational Research Journal.

Graue, E. & Walsh, D.J. (1997). *Studying children in context: Theory, methods, & ethics*.

Thousand Oaks: Sage Publications.

Grissmer, D. (1999). Class size effects: Assessing the evidence, its policy implications, and

future research agenda. Conclusion. *Educational Evaluation and Policy Analysis, 21*(2), 231-248.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Thousand Oaks: Sage

Publications.

Hanushek, E. A. (1999). Some Findings from an Independent Investigation of the Tennessee

STAR Experiment and from Other Investigations of Class Size Effects. *Educational Evaluation and*

Policy Analysis, 21(2), 143-163.

Harris, D. (2006). *Class size and school size: Taking the trade-offs seriously*: Brookings Institute.

NICHD Early Child Care Research Network. (2004). Does class size in first grade relate to

children's academic & social performance or observed classroom processes? *Developmental Psychology,*

40(5), 651-665.

Nye, B., Hedges, L. V., & Konstantopoulos, S. (2001). The long-term effects of small classes in

early grades: Lasting benefits in mathematics achievement at grade 9. *Journal of Experimental Education,*

69(3), 245-257.

Northwest Foundation for Children, Inc.,/ Responsive Classroom/, Retrieved March 14, 2007

from http://www.responsiveclassroom.org/about/aboutrc.html

Rice, J. K. (1999). The impact of class size on instructional strategies and the use of time in high

school mathematics and science courses. *Educational Evaluation and Policy Analysis, 21*(2), 399-414.

Shepard, L.A. (2000). The role of assessment in a learning culture. *Educational Researcher,*

29(7) p. 4-14.

Shepard, L.A. (2005). Formative assessment: Caveat emptor. A paper presented at the ETS

Invitational Conference, *The future of assessment: Shaping teaching & learning*.

Slavin, R. (1989). Achievement effects of substantial reductions to class size. In R. Slavin (Ed.),

School and classroom organization (pp. 247-257). Hillsdale, NJ: Earlbaum.

Smith, M. L., & Glass, G. V. (1980). Meta-analysis of research on class size and its relationship

to attitudes and instruction. *American Educational Research Journal, 17*(4), 419-433.

Smith, P., Molnar, A., & Zahorik, J. (2003). Class-size reduction: A fresh look at the data,

Educational Leadership (Vol. 61, pp. 72-74): Association for Supervision & Curriculum Development.

Stasz, C., & Stecher, B. M. (2002). Before and after class size reduction. In J. D. Finn & M. C.

Wang (Eds.), *Taking small classes one step further* (pp. 19-50). Greenwich, CT: Information Age

Publishing.

Zahorik, J., Halbach, A., Ehrle, K., & Molnar, A. (2003). Teaching practices for smaller classes.

Educational Leadership, 61(1), 75-77.