Plan of Action for Implementing the SCALE Quality Indicator System

Norman L. Webb
Wisconsin Center for Education Research
University of Wisconsin—Madison

Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, April 11, 2005

Contact Information:

Norman L. Webb
Wisconsin Center for Education Research
1025 W. Johnson Street
Madison, Wisconsin 53706

The preparation of this report was supported by a grant from the National Science Foundation to the University of Wisconsin–Madison (EHR 0227016) for a Mathematics & Science Partnership project called the System-wide Change for All Learners and Educators (SCALE) Partnership. At UW–Madison, the SCALE project is housed at the Wisconsin Center for Education Research. The other five partners are the University of Pittsburgh, where the SCALE project is housed within the Learning Resources Development Center’s Institute for Learning; the Los Angeles Unified School District; Denver Public School District, Providence Public School District; and Madison Metropolitan School District. Any opinions, findings, or conclusions are those of the author and do not necessarily reflect the view of the supporting agency.
Plan of Action for Implementing the SCALE Quality Indicator System

Norman L. Webb
Wisconsin Center for Education Research
University of Wisconsin–Madison

The SCALE Quality Indicator System is a work in progress. The purpose of this paper is to describe in some detail the design of the indicator system and to provide samples of some statistics that will likely evolve into indicators. The paper concludes with the acknowledgement of major challenges that exist.

An Indicator System to Monitor Change over Time

The SCALE Quality Indicator System consists of a set of indicators that represent major SCALE attributes as described in the strategic plan and annual reports. An indicator is a statistic that has particular meaning related to a policy or program. For example, the Gross Domestic Product (GDP) is an economic indicator that is used to judge the total output of the United States economy. The GDP is computed by considering a multiplicity of factors, including consumer spending, business and residential investments, government spending, and trade deficit. The SCALE Quality Indicator System (SQIS) is envisioned as a collection of measures that have significance to SCALE as mapped out in its theory of change and as is enunciated by the project as it evolves over time.

The indicators in the SQIS will exist in policy contexts in multiple arenas. The indicators are designed to inform the management needs of the SCALE project and its leaders. As such, some indicators in the system are designed to produce information on the degree to which the four major SCALE goals are being attained. Other indicators are to illuminate if the design strategies are working as planned. Beyond the immediate project needs of SCALE, the quality indicator system will address the reporting requirements of the National Science Foundation, the major funding agency.

The Management Information System (MIS) devised for the reporting needs of NSF have specific requirements that go beyond the SCALE reporting needs. The SQIS has been designed to respond to these external needs and to produce information to comply with NSF’s requests for information in a form that can be aggregated across the Mathematics and Science Partnerships. To meet the specific needs of SCALE and NSF, the SQIS depends heavily on the information collected by and needs of the four partner school districts—Los Angeles Unified School District, Denver Public School District, Madison Metropolitan School District, and Providence Public School District. The indicator system is designed to produce information that will be useful to the four school districts. The districts’ cooperation is critical to the success of an effective indicator system. Thus, the SQIS will produce information in a form that is helpful to the school district and that goes beyond each district’s existing reporting capabilities.
Indicator systems are developed over time usually through an iterative process rather than deductively (de Neufville, 1975; Shavelson, McDonnell, & Oaks, 1991). The SCALE Quality Indicator System also will be developed iteratively by responding to the needs of the three major types of stakeholders—SCALE, NSF, and the four partner school districts. The design of the SQIS described in this paper was developed with input from all three stakeholders. However, it is anticipated that many practical issues will arise in data acquisition and processing that will result in modifications to the prescribed indicator system.

Initially, the indicator system is designed to produce descriptive information in four major areas:

- the number of teachers who have been reached through SCALE activities;
- the change in student achievement over time;
- the nature of instructional practices by teachers; and,
- the participation of students in challenging mathematics and science experiences.

Over time, the indicator system will gather information on schools. The SQIS is not designed to produce variables that can be used in structural equations analyses. As the indicator system matures over time, it may be possible for indicators to be used in more sophisticated analytic modeling. However, the production of descriptive information to meet the needs of the major stakeholders is seen as sufficiently challenging without honing in on the refined variables needed for more analytic modeling.

**Purposes of the SCALE Quality Indicator System**

Three main purposes have guided the design of the SCALE Quality Indicator System. The system has been designed to produce valid and reliable information that can be used to:

1. monitor the progress of the four partner school districts to attain district-defined outcomes related to student achievement, opportunity to learn, instructional quality, and teacher professional learning;
2. judge the degree to which SCALE attains specified benchmarks identified in its strategic plan; and,
3. report information to the National Science Foundation to comply with the Management Information System requirements.

Note that the purpose of the SQIS is not to track all of the activities and policies of the four SCALE school districts. There are too many activities and policies for this to be done with the available resources. Rather, the SQIS has been designed to address the most important information and factors as determined by SCALE’s goals, the requirements imposed on SCALE by NSF, and the functions of the districts influenced by SCALE. It is sufficient to build an indicator system that can describe with some degree of accuracy the changes that occur in the four school districts and to track the SCALE activities that may be associated with these changes.

The SQIS is one of many means used by the SCALE evaluation and research team to study the impact of SCALE on the school districts and on student achievement. The
Building Partnership team and the Case Study team will be gathering qualitative data on SCALE and its activities within the school district. The data produced by these teams will help to describe the complexity of the SCALE initiative and its interactions within the school districts. The SQIS is designed to provide primarily quantitative data to complement information produced by the other two evaluation teams and by the four districts. It is expected that the compilation of information from the combination of studies will produce evidence that can be used to judge the value of SCALE and its capacity to meet its goals.

**Major Components of the SCALE Quality Indicator System**

The SCALE Quality Indicator System consists of six major components, each consisting of a set of variables. These components, along with other components, and their relationships are depicted in Figure 1. Two student outcome components will be tracked, student achievement and student participation as represented by the mathematics and science courses students take. Student participation also will include their engagement with immersion units and other challenging mathematics and science instruction.

Instructional activities employed by teachers are directly related to student participation and achievement. These activities can be influenced by SCALE through teachers and the curriculum. SCALE will provide professional learning experiences for teachers and will work with districts to define curriculum materials that will lead to students learning challenging mathematics and science. Some teachers in the school districts, particularly new teachers who graduated from local universities but also classroom teachers, will receive professional learning experiences provided by universities and supported, at least in part, by SCALE.

School participation—providing leadership, guidance for the necessary instructional activities, and support for professional learning activities among teachers including mentoring, coaching, and other peer interactions—is critical to the SCALE theory of change. It is further recognized in this very simplistic design of the SCALE interventions that, to a major degree, SCALE influence will be mediated through the district leadership. The Case Study team will analyze how this mediation works through many different channels, such as an instructional guidance system, to reach schools and teachers. Portraying the complexity of the SCALE interventions and the school districts is left to the Case Study team. To supplement this work, the SQIS will attend primarily to those factors that are directly related to schools, teachers, and students to produce descriptive information of what changes over time. As such, the six main components of the SQIS will be:

1. Student Achievement
2. Student Participation
3. Instructional Activities
4. Teachers (Background and Knowledge)
5. Teacher Professional Learning
6. Schools
Figure 1

SCALE Quality Indicator System

Input
- SCALE
- District

Capacity
- University
- School
- Teachers
- Curriculum
- Professional Learning Experience

Action
- Instructional Activities

Student Outcomes
- Student Achievement
- Student Participation

Note: Shaded areas represent the major focus of the SCALE Quality Indicator System.
This paper continues with identification of the evaluation questions to be addressed by the SQIS. The evaluation questions were derived from relevant benchmarks for each of the four SCALE goals described in the Strategic Plan (June, 2003). Next, selected variables are described that will produce data that answer the evaluation questions. Some preliminary data on some indicators are presented to illustrate the type of data that are to be reported. The paper concludes with a brief discussion of some of the issues we face in developing the indicator system.

**Evaluation Questions**

A number of evaluation questions were developed on the basis of the SCALE benchmarks listed in the SCALE strategic plan. These evaluation questions are listed below under the four major SCALE goals. The explicit benchmark or benchmarks addressed by each evaluation question is referenced after each evaluation question when appropriate. Some of the evaluation questions are not referenced by an explicit benchmark, but are included because of their overall importance to the SCALE theory of change. A main evaluation question is stated for each goal. Then each general question is explicated by more detailed questions. The SCALE Quality Indicator System is designed to produce information that will answer these evaluation questions.

**Goal 1:**
1. What proportion of students experience deep, conceptually based instruction on core mathematics and science concepts on a continuing basis?

**Teachers**
1.1 What number and proportion of teachers have participated in SCALE–supported professional development or received SCALE support? (No explicit benchmark)

1.2 What proportion of teachers provides students with deep, conceptually based instruction on core mathematics and science concepts as described in the Principles of Learning and by Disciplinary Literacy? (No explicit benchmark)

**Schools**
1.3 What proportion of schools provides a climate of support to teachers and exist as a professional learning community so that students experience mathematics and science instruction as described in the Principles of Learning and Disciplinary Literacy? (Benchmark 2c-3, 2d-1, 2d-2, 2d-3, and 3d-1)

1.4 What proportion of principals engages in continuous improvement planning that incorporates the analysis of student assessment data, instructional activities, and student learning? (Benchmark 3c-2)

1.5 What proportion of principals implements a review of benchmarks and student progress, including quarterly reviews with individual teachers? (Benchmark 3c-3, 3c-4)
1.6 What proportion of schools have been identified as low performing, as defined by the district, in mathematics and science achievement, and what proportion of the low performing schools has received assistance and support from the district? (Benchmark 4e-2, 4e-3)

1.7 What proportion of low-performing schools have shown substantial improvement in mathematics and science achievement each year and, as such, are no longer identified as low performing? (Benchmark 4e-4)

**Student Achievement and Participation**

1.8 What proportion of students is judged to be proficient or above in mathematics as determined by each state and/or district? (Benchmark 4b-1)

1.9 What proportion of students completes an algebra course or its equivalent by at least 9th grade? (Benchmark 1a-6)

1.10 What proportion of 9th and 10th grade students completes courses in science and mathematics that satisfy the high school graduation requirements and qualify them for higher education? (Benchmark 4d-2)

1.11 What proportion of high school students takes college-qualifying mathematics and science courses by the number of courses taken? (No explicit benchmark)

1.12 What are the average mathematics achievement scores by grade by year? (No explicit benchmark)

1.13 What are the average science achievement scores by grade by year (when available)? (No explicit benchmark)

1.14 What is the annual growth of student achievement in mathematics and science by grade and year? (No explicit benchmark)

**Goal 2**

2. What proportion of teachers of science and mathematics (K–12) has engaged students annually in an extended (e.g., four-week) scientific investigation (i.e., immersion unit)?

2.1 What proportion of students has participated in an immersion-unit experience? (Benchmark 6a-5, 6a-6, 6a-7)

2.2 What proportion of mathematics and science teachers has been trained to implement an immersion unit? (No explicit benchmark)

**Goal 3**

3. What number of inservice teachers has participated in STEM–IHE-led development programs that give teachers a deeper grasp of STEM content and effective pedagogical strategies for engaging students in learning?
3.1 What number of teachers has participated in content courses taught or research experiences organized by STEM faculty during the year? (New Benchmark in Year 3 Implementation Plan Track 3-1, Track 3-2, Track 3-4)

**Goal 4**

4. To what degree has education of all students in the school districts become more equitable in that minority and female students enroll in and complete mathematics and science courses at a comparable rate as White male students and with a comparable level of achievement?

4.1 What proportion of minority and female students enrolls in and completes college-qualifying mathematics and science high school courses? (Strategic Plan Goal 4)

4.2 What proportion of minority and female students completes high school, enrolls in college, and pursues fields in mathematics and science? (Strategic Plan Goal 4)

4.3 To what degree have the equity gaps closed? (10b-1 and 2, Year 3 Implementation Plan) (Note: More detailed benchmarks for Goal 4 are under development.)

**Critical Variables**

To answer the evaluation questions for SCALE and to produce the requested information for the MIS by NSF, a series of general variables have been identified. These general variables are organized by the six main components of the indicator system. The information for each variable will be needed by school year, grade, gender, race, poverty, special education status, and English language learner status. The numbers in parentheses refer to the evaluation questions above that are related to the variable.

**Achievement**

- Individual student mathematics proficiency ranking on an annual test by school (1.8)
- Individual student science proficiency ranking on an annual test by school (1.8)
- Individual student mathematics achievement score on an annual test by school (1.12)
- Individual student science achievement score on an annual test by school (1.13)
- Individual student annual growth in mathematics achievement by school (grades 3-4, 4-5, 5-6, 6-7, and 7-8) (1.14)
- Individual student annual growth in science achievement for grades assessed by school (1.14)
- Degree to which mathematics achievement scores on an annual test have achieved parity by male/female and race (4.3)
- Degree to which science achievement scores on an annual test have achieved parity by male/female and race (4.3)

**Student Participation**

- Total number of students
- Number of 10th graders who have completed 9th grade algebra (1.9)
Number of 9th and 10th graders enrolled in high school graduation qualifying mathematics courses (1.10)
Number of 9th and 10th graders enrolled in high school graduation qualifying science courses by year (1.10)
Number of 12th grade students by the number of college-qualifying mathematics semesters completed (1.11)
Number of 12th grade students by the number of college-qualifying science semesters completed (1.11)
Number of 12th grade students who have completed college-qualifying mathematics courses by course title (Algebra I, Geometry, Algebra II, etc.) (1.11, 4.1)
Number of 12th grade students who have completed college-qualifying science courses by course title (Biology I, Chemistry I, etc.) (4.1)
Number of graduating 12th grade students who have completed mathematics course requirements for admission to state higher education institutions (1.10)
Number of graduating 12th grade students who have completed science course requirements for admission to state higher education institutions (1.10)
Number of female and minority students who attend higher education after high school and major in mathematics- and science-related fields (4.2)
Number of students who have engaged in an immersion unit or extended scientific (mathematical) investigation during the school year (4.2)
Number of students by grade who participated in an immersion unit or an extended scientific investigation during the school year (2.1)

Teacher Background
Total number of teachers of mathematics (K–5 or 6 all and 6 or 7–12 content teachers)
Total number of teachers of science (K–5 or 6 all and 6 or 7–12 content teachers)
Grade range teaching for current school year
Type of certification
Content area of certification (mathematics, science, elementary, K–8, secondary, other)
Undergraduate institution issuing undergraduate degree
1. Undergraduate major (mathematics, science, education, other)
2. Undergraduate minor (mathematics, science, education, other)
Type of any advanced degrees (mathematics, science, education, other)
Number of years teaching in general and in the district
Number of years teaching mathematics
Number of years teaching science
Teaching mathematics or science and other subjects (self-contained classroom), only one of the two, or teach mathematics/science part time
School level (elementary, middle, junior high, K–8, high school 9–12, high school 10–12, etc. Note: LAUSD has about 15 levels)
Amount of school year taught at school (full, date arrived after beginning of school year, date left before school year) (teacher attrition)
Number of minutes of mathematics taught per week
Number of minutes of science taught per week
Mathematics courses taught during school year (by course IDs)

1
2
Elementary school mathematics courses
Middle school mathematics courses (school with grade 8)
High school mathematics courses (grades 9–12)

Science courses taught during school year
Elementary school science courses
Middle grades science courses (school with grade 8)
High school science courses (grades 9–12)

**Teacher Professional Learning**
- Number of teachers of mathematics who have received training on implementing immersion units or other extended scientific investigation (2.2)
- Number of teachers of science who have received training on implementing immersion units or other extended scientific investigation (2.2)
- Number of teachers of mathematics who have received STEM IHE-led professional development programs on STEM content and pedagogical strategies during the year and prior to the school year (G3, 3.1)
- Number of teachers of science who have received STEM IHE-led professional development programs on STEM content and pedagogical strategies during the year and prior to the school year (G3, 3.1)
- Number of teachers of mathematics who have participated in a research experience organized by STEM faculty/staff during the year (3.1)
- Number of teachers of science who have participated in a research experience organized by STEM faculty/staff during the year (3.1)

**Instruction**
- Number of teachers who provide deep, conceptually based instruction on core mathematics and science concepts (G1, 1.1)
- Number of teachers who implement the Principles of Learning in mathematics
- Number of teachers who implement the Principles of Learning in science
- Number of teachers of mathematics who have engaged students in an extended scientific investigation during the school year (G2)
- Number of teachers of science who have engaged students in an extended scientific investigation during the school year (G2)

**School**
- Number of schools by grade range
- Number of schools by levels of professional learning community and supportive climate for deep, conceptually based instruction (1.2)
- Number of schools with a principal who supports continuous improvement among teachers, analysis of student assessments, and teachers’ quarterly review of benchmarks and student progress (1.3, 1.4)
- Number of schools in a district identified by the district as low performing (1.5)
- Number of schools previously identified low-performing schools that have substantial improvement in student achieve for the year (1.5)

---

G3 refers to SCALE goal 3. This specific variable is related to the general goal, in addition to any specific benchmarks under the general goal.
Number of schools previously identified as low performing schools that were removed from this status during the school year (1.6)
Number of schools identified as low performing that received assistance and support from the district during the school year (1.5)

Sample Indicators

We have received baseline data from all four SCALE school districts. The data received varies by year and variables. Madison Metropolitan School District supplied us with four years of data from 1999 through 2002. For each of these years, the State of Wisconsin administered the Wisconsin Knowledge and Concept Examination (WKCE), a version of the TerraNova, to students in grades 4, 8, and 10. During those years, the tests were administered in February. Beginning in 2003, the administration time was changed to October. We received both scale scores and level-of-proficiency scores for all students who were tested for each of the grades. One indicator we will report will be change in scale scores over time. A complication we face in doing this meaningfully will be to account for changes in the time of testing, such as Wisconsin’s change in 2003 from February testing to October testing, and a change in tests, such as California’s change in 2003 from the SAT9 to the CAT6.

Another indicator we will report is related to the percent proficient. Change in the percent proficient can be more difficult to interpret than change in scale scores because states and districts can modify the cutoff score required for a student to be designated as proficient. However, the statistic does have importance due to provisions of the No Child Left Behind Act, which currently require states to report on the percent of students who are proficient in mathematics and reading and will hold states and schools accountable for the percent of students who are proficient. One of the SCALE Goal 1 benchmarks is to target a 5% increase at each proficiency level per year K–8 (Benchmark 4b-1 of the SCALE Strategic Plan). Because of this benchmark and the prominence given to percent of students who are proficient at the national level, one of the indicators we have developed is the rate of change in percent proficient disaggregated by demographic groups. Table 1 shows one means we have chosen to report the information for one of the four school districts, Madison. In the baseline years, prior to SCALE’s implementation in mathematics, the total group of students for the three grade levels had little, or a slightly negative, rate of change in the percent proficient over the four baseline years. However, some Black students had small positive rates of change at all grade levels. In science, the rate of change was negative for the three grade levels.
Table 1  
*Rate of Change in Percent Proficient over Four Baseline Years—February 1999, 2000, 2001, and 2002—For Mathematics and Science*  
**Madison Metropolitan School District**

<table>
<thead>
<tr>
<th>Grade</th>
<th>Mathematics Percent proficient in last baseline year</th>
<th>Mathematics Range of Percent proficient across baseline</th>
<th>Mathematics Change per year</th>
<th>Science Percent proficient in last baseline year</th>
<th>Science Range of Percent proficient across baseline</th>
<th>Science Change per year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 4</td>
<td>63</td>
<td>63-70</td>
<td>-2.80</td>
<td>69</td>
<td>81-69</td>
<td>-5.10</td>
</tr>
<tr>
<td>Am. Indian</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asian</td>
<td>61</td>
<td>52-61</td>
<td>3.30</td>
<td>57</td>
<td>57-62</td>
<td>-1.10</td>
</tr>
<tr>
<td>Hispanic</td>
<td>45</td>
<td>36-50</td>
<td>-2.00</td>
<td>42</td>
<td>42-67</td>
<td>-9.70</td>
</tr>
<tr>
<td>White</td>
<td>78</td>
<td>78-83</td>
<td>-2.00</td>
<td>85</td>
<td>81-91</td>
<td>-2.50</td>
</tr>
<tr>
<td>Grade 8</td>
<td>48</td>
<td>45-59</td>
<td>-0.10</td>
<td>58</td>
<td>58-65</td>
<td>-1.60</td>
</tr>
<tr>
<td>Am. Indian</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asian</td>
<td>41</td>
<td>35-41</td>
<td>1.60</td>
<td>48</td>
<td>43-50</td>
<td>-0.40</td>
</tr>
<tr>
<td>Black</td>
<td>13</td>
<td>8-13</td>
<td>0.60</td>
<td>20</td>
<td>20-26</td>
<td>-1.10</td>
</tr>
<tr>
<td>Hispanic</td>
<td>25</td>
<td>21-29</td>
<td>-0.80</td>
<td>27</td>
<td>27-50</td>
<td>-5.80</td>
</tr>
<tr>
<td>White</td>
<td>62</td>
<td>58-62</td>
<td>-0.70</td>
<td>72</td>
<td>72-77</td>
<td>-0.50</td>
</tr>
<tr>
<td>Grade 10</td>
<td>45</td>
<td>45-53</td>
<td>-0.20</td>
<td>48</td>
<td>48-56</td>
<td>-0.80</td>
</tr>
<tr>
<td>Am. Indian</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asian</td>
<td>33</td>
<td>32-44</td>
<td>-0.90</td>
<td>31</td>
<td>31-38</td>
<td>-0.80</td>
</tr>
<tr>
<td>Black</td>
<td>12</td>
<td>12-16</td>
<td>0.00</td>
<td>16</td>
<td>15-19</td>
<td>0.60</td>
</tr>
<tr>
<td>Hispanic</td>
<td>23</td>
<td>20-25</td>
<td>0.80</td>
<td>24</td>
<td>20-33</td>
<td>0.70</td>
</tr>
<tr>
<td>White</td>
<td>58</td>
<td>56-65</td>
<td>1.10</td>
<td>61</td>
<td>61-68</td>
<td>-0.30</td>
</tr>
</tbody>
</table>
A very important goal for SCALE is to have every student experience deep, conceptually based instruction on core mathematics and science concepts on a continuing basis. We are in the process of developing some indicators that can be used to monitor the attainment of this goal. Because the goal is to have every student experience deep, conceptually based instruction, we are limited to using self-report information from teachers collected by written questionnaires. Surveying all teachers, or a sufficiently large random sample of teachers, presents a number of challenges, including getting the cooperation of the school districts and teachers. We have had some success in working cooperatively with staff in two of the districts to develop and administer surveys that

Table 2

Mean, Median, and Standard Deviation for Inquiry-based Learning Questions and Scale (N=68)

Denver Public Schools

<table>
<thead>
<tr>
<th>Question</th>
<th>Mean</th>
<th>Median</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How often have lessons taught in the past school year focused on addressing specific components of the state science standards?</td>
<td>3.99</td>
<td>4</td>
<td>1.169</td>
</tr>
<tr>
<td>2. How often do students in a typical class explain their reasoning, whether written or verbal?</td>
<td>3.41</td>
<td>3</td>
<td>.909</td>
</tr>
<tr>
<td>3. How often do students in a typical class support their point of view based on sound reasoning and facts?</td>
<td>3.13</td>
<td>3</td>
<td>1.02</td>
</tr>
<tr>
<td>4. How often have lessons taught in the past school year used an overall theme that continued throughout the school year?</td>
<td>3.11</td>
<td>3</td>
<td>1.518</td>
</tr>
<tr>
<td>5. How often have lessons taught in the past school year focused on studying a topic in depth, rather than covering basic facts or procedures?</td>
<td>2.96</td>
<td>3</td>
<td>1.3</td>
</tr>
<tr>
<td>6. On average, how often do you facilitate whole-class discussions in which you talk less than the students?</td>
<td>2.86</td>
<td>3</td>
<td>1.04</td>
</tr>
<tr>
<td>7. On average, how often do students examine a particular topic or issue in depth, beyond basic facts and summarizations?</td>
<td>2.8</td>
<td>3</td>
<td>1.001</td>
</tr>
<tr>
<td>8. How often have lessons taught in the past school year required students to organize, interpret, and use information to produce a piece of original work?</td>
<td>2.72</td>
<td>3</td>
<td>.958</td>
</tr>
<tr>
<td>9. How often do students in a typical class collect data or conduct an experiment to answer a question posed by the textbook or teacher?</td>
<td>2.6</td>
<td>3</td>
<td>.769</td>
</tr>
<tr>
<td>10. How often have lessons taught in the past school year required students to use science concepts and information in settings different than the one it is learned?</td>
<td>2.57</td>
<td>3</td>
<td>.941</td>
</tr>
<tr>
<td>11. How often do students in a typical class participate in student-led discussions?</td>
<td>2.4</td>
<td>2</td>
<td>1.197</td>
</tr>
<tr>
<td>12. How often do students in a typical class collect original data related to a course topic?</td>
<td>2.26</td>
<td>2</td>
<td>.896</td>
</tr>
<tr>
<td>13. How often do students in a typical class generate an original scientific question related to a course topic?</td>
<td>2.19</td>
<td>2</td>
<td>1.146</td>
</tr>
<tr>
<td>14. On average, how often do you coordinate the content of your courses with other teachers?</td>
<td>2.17</td>
<td>2</td>
<td>1.351</td>
</tr>
<tr>
<td>15. How often do students in a typical class collect data to answer a self-generated scientific question?</td>
<td>1.99</td>
<td>2</td>
<td>.931</td>
</tr>
</tbody>
</table>

Scale: mean = 2.74, median = 2.73, standard deviation = .607

r = .84

5-point scale: Rarely/Never (1), At Least Once a Month (2), At Least Once a Week (3), 2-3 Times a Week (4), Almost Daily (5)

Note: Shaded items are the same or similar to items used for Madison (see Table 3).
produce information useful to the districts on science instruction, while serving the needs of SCALE. We are using scales formed by aggregating data across a number of items to create an indicator related to different instructional approaches. One indicator we have developed is called Inquiry-Based Learning. Table 2 shows the 15 Likert-scale items that are used to form a version of this scale and the results of administering a questionnaire to 68 middle school and high school science teacher leaders in Denver. Table 3 shows another version of the scale of items (n = 14 items) administered to elementary through grade 8 teachers of science in Madison. Although some items are the same for both administrations, most items varied. A total of 281 elementary teachers (46% of the total) and 52 middle school science teachers (68% of the total) returned completed questionnaires. In the spirit of partnership and trying to produce information that is responsive to district needs, we developed the scale we used in each district with the input of district science leadership. As a result, the set of items administered in Madison varied some from those administered in Denver to respond to perceived district differences. This implies that the indicator system cannot be used to do cross-district comparisons. This is not a problem, since comparing districts is not a purpose of the indicator system.

In both tables, the items are ordered by mean score. In Madison, about 75% of teachers who responded had inquiry based-scale scores between 2.5 and 5, indicating that they used inquiry-based learning once-a-week-to-almost-daily. In Denver (Table 2), respondents’ overall scores on the inquiry-based learning scale ranged from 1.6 to 4.07, signifying use of these classroom practices and activities roughly between “At Least Once a Month” to “2-3 Times a Week” for all teachers. Thirty-percent of the teachers reported using inquiry-based practices regularly (“Once a Week” to “2-3 Times a Week”) in their science classrooms.

The Inquiry-based Learning Scale has the potential to be converted into an indicator of instructional practices. However, we are still in the process of developing a scale that will represent the critical foci of the project. The Principles of Learning developed by the Institute for Learning is an important driver of interventions to influence instructional practices in the four districts. Principles of Learning incorporate some of what is considered inquiry-based learning, but is more explicit and more comprehensive in clarifying what quality instruction is and the context for quality instruction. We are still in the process of determining what items to include, what Likert scale to use, and how best to report the information.
Table 3
Mean, Median, and Standard Deviation for Inquiry-based Learning Questions and Scale (N = 333)
Madison Metropolitan School District

<table>
<thead>
<tr>
<th>Question:</th>
<th>Mean</th>
<th>Median</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. How often did students in a typical class last year explain their reasoning, whether written or verbal?</td>
<td>3.42</td>
<td>3</td>
<td>.990</td>
</tr>
<tr>
<td>4'. How often did lessons you taught last school year use themes/big ideas about specific content areas?</td>
<td>3.37</td>
<td>4</td>
<td>.867</td>
</tr>
<tr>
<td>5'. How often did lessons you taught last school year focus on studying a topic in depth, rather than covering basic facts or procedures?</td>
<td>3.28</td>
<td>3</td>
<td>1.056</td>
</tr>
<tr>
<td>20. How often did lessons you taught last school year allow students time to reflect on their work?</td>
<td>3.24</td>
<td>3</td>
<td>.968</td>
</tr>
<tr>
<td>17. How often did lessons you taught last school year have students explain their reasoning to you or to their classmates how the topic related to their personal experiences or to a real world problem?</td>
<td>3.10</td>
<td>3</td>
<td>1.013</td>
</tr>
<tr>
<td>6'. How often did you facilitate whole-class discussions in which you talked less than students in the science classes you taught last year?</td>
<td>3.08</td>
<td>3</td>
<td>1.086</td>
</tr>
<tr>
<td>15'. How often did students in a typical class last year collect data to answer a scientific question?</td>
<td>2.94</td>
<td>3</td>
<td>1.001</td>
</tr>
<tr>
<td>8. How often did lessons you taught last school year require students to organize, interpret, and use information to produce a piece of original work?</td>
<td>2.89</td>
<td>3</td>
<td>1.086</td>
</tr>
<tr>
<td>11. How often did students in a typical class last year participate in student-led discussions?</td>
<td>2.73</td>
<td>3</td>
<td>1.175</td>
</tr>
<tr>
<td>7'. How often did you examine a particular topic or issue in depth, beyond basic facts and summarizations in the science classes you taught last year?</td>
<td>2.71</td>
<td>3</td>
<td>1.073</td>
</tr>
<tr>
<td>29. How often did students in a typical class last year prepare written science reports, assignments, or reflections?</td>
<td>2.24</td>
<td>2</td>
<td>1.076</td>
</tr>
<tr>
<td>21. How often did lessons you taught last school year allow students the opportunity to investigate and analyze a topic of their choosing?</td>
<td>2.21</td>
<td>2</td>
<td>1.195</td>
</tr>
<tr>
<td>13. How often did students in a typical class last year generate an original scientific question related to a course topic?</td>
<td>2.20</td>
<td>2</td>
<td>.988</td>
</tr>
<tr>
<td>22. How often did you assign projects or investigations of at least one week’s duration in the science classes you taught last year?</td>
<td>2.04</td>
<td>2</td>
<td>1.013</td>
</tr>
</tbody>
</table>

5-point scale: Rarely/Never (1), At Least Once a Month (2), At Least Once a Week (3), 2-3 Times a Week (4), Almost Daily (5)

Note: Shaded items are the same or similar (indicated by ‘’) to items used for Denver (see Table 2).

Challenges

The SCALE Quality Indicator System is still in a very formative stage. Major work now is to develop the instruments, procedures, and cooperation to gather the needed data. Three questionnaires are under development:

Teacher Background and Professional Development
Principles of Learning
School Climate
We have collected baseline information on student achievement from all four districts and are in the process of obtaining student achievement data for the first year of SCALE.

One major challenge we face is how to aggregate data in a meaningful way to report statistics that can monitor change over time while representing the theory of change in SCALE’s intervention. A second challenge is to assure that we have valid indicators. Teacher report data on instructional practices can easily be tainted: What teachers report they do is not always what teachers actually do in their classrooms. But we are restricted to using questionnaires mainly because of the size of the school districts and our hope to describe in some way the range in instructional practices in mathematics and science for an entire district. We intend to validate our instruments by correlating the results with a few teachers with classroom observations.

Until now, we have worked cooperatively with district staff to customize scales on instructional practices to better meet the interests and data collection needs of a district. This is very time consuming. Our challenge is to be as responsive as possible to the districts’ needs, while employing a system that is manageable with the provided resources. While the data we have gathered to this point and the development process have been informative, in the future the scales we will develop will need to more applicable across the four districts.

References
