



**ALBUQUERQUE
PUBLIC SCHOOLS**

**Continuous Classroom Improvement
Program
Evaluation Study**

Spring 2008

District Goal: Academic Excellence

May 2008
Debra Heath



ALBUQUERQUE PUBLIC SCHOOLS

BOARD OF EDUCATION

MARY LEE MARTIN
President

BERNA FACIO
Vice President

DOLORES GRIEGO
Secretary

GORDON ROWE
Instructional Policy Chair

PAULA MAES
District Relations Chair

ROBERT D. LUCERO
Finance/Audit Chair

MARTIN R. ESQUIVEL
Capital Outlay Chair

LINDA SINK
Interim Superintendent

RAQUEL REEDY
Associate Superintendent

EDUARDO B. SOTO
Associate Superintendent

THOMAS SAVAGE
Deputy Superintendent

RESEARCH, DEVELOPMENT AND ACCOUNTABILITY

930-A Oak Street SE
Albuquerque, New Mexico 87106
(505) 848-8710
www.rda.aps.edu
Rose-Ann McKernan
Executive Director
Instructional Accountability

INTRODUCTION

Regional Quality Center

In August 2005, APS established a Regional Quality Center (RQC) to continue work conducted by the previously named Office of the Assistant Superintendent for Superintendent's Schools. The mission of the Regional Quality Center is to support continuous improvement in APS using a Baldrige-based systems approach. The center is directed by the APS Assistant Superintendent for Continuous Improvement and staffed by two professional staff and two office assistants. Funding for the RQC comes from the New Mexico Public Education Department (PED) and the Governor's Business Executives for Education.

Continuous Classroom Improvement Program

The Continuous Classroom Improvement program is one of RQC's main initiatives. Continuous Classroom Improvement (CCI) is a model for deploying Baldrige educational improvement strategies in the classroom. Teachers and students use the *Plan Do Study Act* process to problem-solve, set weekly learning goals, develop classroom and student action plans, track progress, and evaluate what worked and plan improvements. The CCI program's ultimate goal is to improve student performance.

In the first two days of school:

- Teachers introduce and post the year's 'strategic learning goal.' For example: *All first graders will read at or above the second grade reading level by the end of the first grade (tested by the DRA two times a year).*
- Teachers present students with the standards they will be expected to master. For example: *I can match words I hear with words on a page.*
- Teachers and students develop a mission statement to guide their work. For example: *We will come to school every day and be the best we can be. We will learn math, English, writing, science, reading and social studies. We will do our daily work and turn our homework in. We will pay attention and respect others while we are learning.*
- Teachers post baseline data showing entering class performance. For example, a graph might show how many students are proficient in each of the kindergarten standards.

On a weekly basis, teachers lead students through the following processes:

- **Plan:** Teachers present and post the week's learning goal, with the way performance will be measured and what the criteria for success will be. For example: *All students will be able to demonstrate different ways to count the total number of squares in arrays. We will take a quiz on Wednesday that lets us demonstrate our understanding. A score of 80% or better out of 100 means we've got it!*

- **Do:** The class plans and implements strategies and activities to meet the weekly learning goal. For example: *The teacher will give examples. Students will underline key words and practice with a buddy.*
- **Study:** The class measures how well they learned the topic, usually by charting results of tests and quizzes. Teachers and students analyze what facilitated learning and what changes would enhance learning, often via the *Plus/Delta* technique.
- **Act:** The class discusses what students *and* teachers can do to improve for the next learning target. For example: *The teacher will keep explanations short and simple. Students will help develop a Buddy Practice Time rubric to make sure that we don't waste our practice time.*

RQC originally offered CCI training and technical assistance to schools on a voluntary basis and with a gradual implementation timeline. RQC asked each participating school to send two teachers and an instructional coach to CCI training sessions. The goal was that all participating teachers would develop CCI demonstration classrooms, which would serve as models and catalysts for CCI adoption by other teachers. In the winter of 2006-07, PED mandated school-wide CCI deployment in all Priority Schools.

In the program's first year, 2006-07, 30 elementary and middle schools participated, plus support staff of the Highland Cluster. Most participating schools had Restructuring (R1 or R2), Corrective Action or School Improvement designations. In 2007-08, all but two schools continued in the CCI program, for a total of 28 public schools plus the 13 schools of the Highland Cluster. Also in 2007-08, one charter school joined and four schools became "CCI demonstration schools," committing to move toward CCI implementation in all classrooms.

Evaluation Study Purpose & Questions

RQC asked APS' Research, Development and Accountability department (RDA) to help evaluate the effectiveness of its CCI program. After initial conversations, RDA proposed a preliminary study to identify possible process improvements, and to establish whether, when and how an outcome evaluation should be conducted. The evaluation questions are:

1. What is the CCI program's *theory of change*? What outcomes does APS' CCI program hope to produce and through what activities and processes?
2. How does the CCI program's *theory of change* align to research?
3. What do implementation experiences to-date tell us about the strengths and weaknesses of the CCI program?
4. What programmatic enhancements, if any, could increase the CCI program's chances of achieving intended outcomes?
5. What design elements, methods, resources and timeline are recommended for evaluating CCI program outcomes?

Evaluation Study Methods

The RDA evaluator met with RQC staff multiple times to develop a clear understanding of the CCI program's purposes, plans, activities, and experiences to-date. RQC staff shared documents that summarize program plans, participants' CCI proficiency levels, and participant feedback. In addition, the RDA evaluator reviewed documents created by Jim Shipley & Associates, CCI-related documents available on the web, and research literature. Teacher-level data collection included a school site visit with classroom observations of PDSA practices, as well as a brief survey of formative assessment practices conducted at a meeting of 'proficient' CCI teachers in spring 2008.

RESULTS

CCI Program Logic and Theory of Change

Evaluation Question #1: What is the CCI program's theory of change? What outcomes does APS' CCI program hope to produce and through what activities and processes?

RDA created a logic model to visually depict CCI activities and their intended short, medium and long-term outcomes (Figure 1). The logic model shows the cause and effect relationships that the RQC envisions to achieve its ultimate goal. This is the CCI program's *theory of change*.

The goal provided by the PED in 2007-08 is for CCI to improve reading and/or math student achievement in RQC demonstration schools and/or classrooms, as measured by state and district assessments. RDA suggests a more specific and measurable goal: to increase students' rate of learning in reading and math. This is depicted in Figure 1, item 6.0.

Starting from the far left side of the logic model (Figure 1), RQC planned and conducted the following activities to promote and disseminate high quality CCI practices:

- Train, coach and provide rewards and recognition to CCI demonstration teachers and instructional coaches (1.1) so they can train and coach others in their schools (1.2).
 - In Year One (2006-07), train two teachers in each school to implement CCI on a demonstration basis. RQC asked schools to select highly experienced teachers who showed an active interest in CCI implementation.
 - Train instructional coaches (ICs), as available, to support demonstration classrooms, model CCI strategies, and train additional teachers in the school (Year One).
 - Conduct monthly site visits to each school to observe classroom and IC practices, coach teachers, consult with students, and collect self-assessment measures of CCI proficiency (Year One).

- Convene meetings of participating teachers four times in Year One.
- In Year Two (2007-08), meet with groups of proficient teachers and ICs at each school two times, to support and expand their implementation of CCI, and to support their efforts to train additional teachers in CCI methods.
- Meet with groups of non-proficient, ‘progressing’ teachers and ICs three times, visit their classrooms four times, and provide a half-day substitute for each teacher to visit another school’s proficient CCI classroom (Year Two).
- Help teachers and school leaders develop CCI deployment plans that specify the targets, timeline and processes for training, implementing, supporting and diffusing CCI (1.3).
- Provide school leaders with training in developing Educational Plans for Student Success (EPSS) that incorporate CCI goals and processes, and conduct school needs assessments, staff presentations and staff professional development (1.4).

RQC expects that as a result of these activities, teacher knowledge of CCI techniques and self-efficacy and commitment to employ them will increase (2.1). Each school should have a detailed CCI deployment plan (2.3). In addition, school leaders should gain increased knowledge of CCI and should demonstrate support for CCI (2.4). For example, leaders may provide teachers with time to receive training and coaching, and may publicly recognize teachers’ CCI accomplishments.

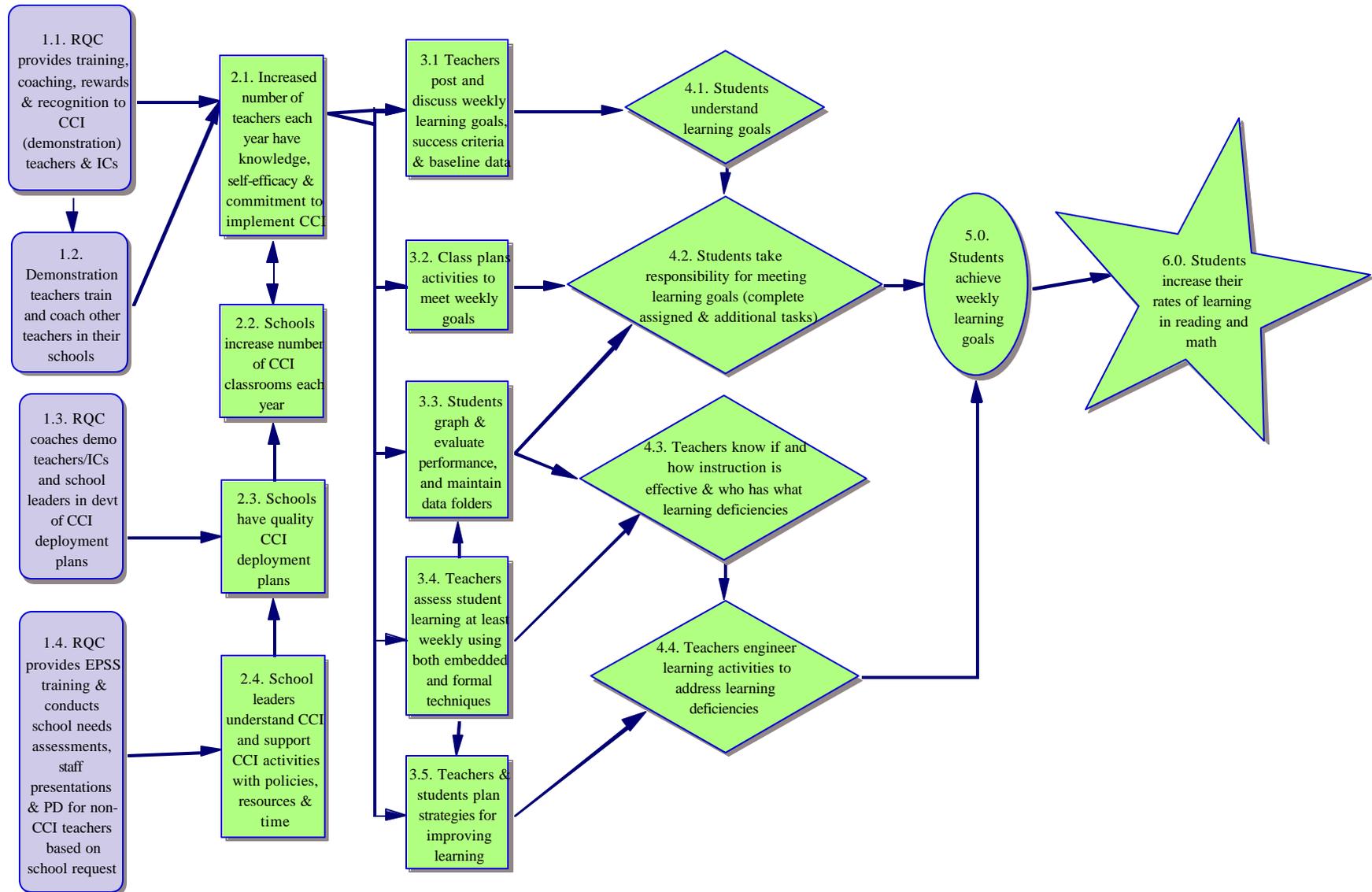
RQC expects that each year the number of teachers implementing CCI proficiently will increase. As CCI diffuses to more classrooms (2.2), CCI will reach more students, and increasing numbers of students will experience CCI for multiple years, enhancing the likelihood of measurable impact.

CCI’s main strategy to increase student learning is the Plan Do Study Act (PDSA) weekly cycle, depicted in logic model items 3.1 to 3.5. According to RQC staff, PDSA processes:

- Help students understand and take responsibility for meeting individual and classroom learning goals (4.1 and 4.2); and
- Help teachers know, on an ongoing basis, if their methods are effective and which students need extra help (4.3).

Teachers who use this information to engineer learning activities that address learning deficiencies in an ongoing way (4.4) are able to help students meet weekly learning targets (5.0). These targets should be calibrated to help students move more quickly and reliably toward and beyond mastery of grade-level standards in reading and math (6.0).

Figure 1. Continuous Classroom Improvement Program Logic Model/Theory of Change



Alignment to Research

Evaluation Question #2: How does the CCI program's theory of change align to research?

RDA found no published research or publicly available descriptions of studies or evaluations of Continuous Classroom Improvement and/or other Baldrige practices and their effects on student outcomes. However, research, theories and models from the fields of education, communication, and psychology provide important insights into likely strengths of the RQC's CCI program. They also highlight opportunities for RQC to enhance the CCI program for maximum success.

Formative Assessment

The literature on formative assessment may provide the most compelling evidence in favor of Continuous Classroom Improvement. Formative assessment is the use of assessment processes, materials and results to maximize student learning during the course of instruction and to inform future instruction.¹ In other words, formative assessment helps 'keep learning on track' toward desired outcomes.

Research shows that formative assessment is a powerful tool for increasing student achievement. In their widely-cited 1998 research synthesis, Black and Wiliam report that formative assessment produces significant learning gains, with effect sizes between 0.4 and 0.7. At the low end (effect size of 0.4), this means raising the average intervention student's achievement to the level of the top 35% of non-intervention students. Moreover, Black and Wiliam's review shows that the largest gains occur for the lowest achievers.² Subsequent studies have concluded that formative assessment is more cost-effective than other popular educational initiatives including class-size reduction, increases in teacher content knowledge, increased per-pupil expenditures, and increased accountability.³

A wide range of practices and products use the term 'formative assessment,' but few of them actually embody the characteristics shown to improve student learning.⁴ Effective formative assessment requires more than assessing frequently and using results for instructional improvement. Researchers have found that the greatest gains in student achievement occur when students clearly understand the learning targets, are actively engaged in assessing their own progress toward those targets, receive descriptive feedback from teachers and peers, and receive guidance about exactly what steps they can take to improve.⁵

Table 1 lists key formative assessment strategies and illustrates their application within the CCI model.⁶ It is clear that CCI embodies many effective formative assessment principles. Even at its most basic level of application, CCI promotes teachers' practice of formative assessment; this is shown in the second column of Table 1, called Basic CCI Applications. Enhanced Applications illustrate what CCI may look like at its highest level of practice, and/or represent opportunities to expand upon formative assessment basics.

Table 1. Alignment of CCI to Formative Assessment Strategies

Formative Assessment Strategies	Basic CCI Applications	Enhanced Applications
Clarify learning goals and criteria for success	Post & discuss mission, standards, learning goals, baseline data & how student performance will be measured.	Have students review samples of work at different standards & discuss evidence. Review scoring rubrics.
Conduct frequent assessments	Conduct assessments weekly.	Use informal, embedded techniques to assess student learning on a daily or more frequent basis.
Engineer effective classroom discussions, questions, activities, and tasks that elicit evidence of student learning	Use state, local & classroom tests or quizzes and demonstrations to measure student proficiency. Some teachers may use embedded assessment techniques.	Use planned questions, classroom discussions & techniques like <i>Thumbs Up</i> , <i>Thumbs Down</i> to check student understanding. Use benchmark tests formatively.
Provide feedback that moves learning forward	Have students review test scores and grades, determine whether they met goals, identify what worked and did not work, & plan strategies for improvement.	Use written and oral comments and questions to help students improve. Minimize emphasis on grades.
Activate students to take ownership of their own learning	Use PDSA process to have students plan increasingly better ways to meet classroom and individual learning goals.	Conduct quick self-assessments like <i>Traffic Lights</i> before, during & after instruction.
Activate students as learning resources for one another	Use classroom PDSA process to have students plan ways they can help each other learn.	Use written and verbal peer assessment techniques. Activate peer discussions after self-assessments.

Clarify Learning Goals and Criteria for Success

Effective formative assessment begins by clarifying for students what they need to learn and at what level. CCI provides teachers with a specific format for writing learning goals. Teachers translate grade-level standards into weekly learning goals with specific performance criteria, all written in student-friendly language. For example: *All students will be able to demonstrate different ways to count the total number of squares in arrays. We will take a quiz on Wednesday that lets us demonstrate our understanding. A score of 80% or better out of 100 means we've got it!*

At an enhanced level of formative assessment practice, teachers help students understand the specific qualities that represent ‘good’ work.⁷ Specifying a target of ‘80% correct’ is not enough. Displaying pictures, examples and/or descriptions of proficient work, as some CCI classrooms do, goes much farther. Another proven strategy is to give students

time to talk about what would count as quality work, and giving students examples of work at different standards to review and discuss. This strategy has been shown to cut the achievement gap between the highest and lowest achieving students in half while raising the performance of the entire class to top levels.⁸ Sharing and creating scoring rubrics with students is another strategy for clarifying performance criteria that proficient CCI teachers report.

Conduct Frequent Assessments

Research shows that effective formative assessment is frequent (minute-by-minute, day-by-day, weekly and monthly).⁹ As Popham (2006) notes, assessment results must be available within a class period or in the midst of an instructional unit. If the results don't get back in time for teachers to adjust instruction on the current topic, he says, then it's not formative assessment.¹⁰

CCI promotes a weekly cycle of assessment, data analysis and instructional fine-tuning, within which more frequent assessments may occur. Data collected from proficient CCI teachers in spring 2008 suggest that most conduct weekly or more frequent assessments in at least one content area (see Appendix A). Their reports also suggest that CCI increases the frequency with which teachers assess student learning, as well as the level and speed with which results are used.

Elicit Evidence of Learning

The centerpiece of formative assessment, according to Wiliam (2007), is engineering effective classroom discussions, questions, activities and tasks that elicit evidence of student learning. Assessing student learning *during* instruction, so that confusion can be addressed immediately, is much more efficient than attempting to address confusions discovered later through formal assessments. This is one rationale Wiliam provides for employing embedded assessment techniques, such as *Traffic Lights* and planned questions. Stiggins argues further that assessment results should be used immediately to remedy learning deficiencies so that subsequent efforts meet with success. This addresses the motivational aspect of learning.¹¹

Data collected in spring 2008 from proficient CCI teachers suggest that CCI has expanded the range of tools and techniques participating teachers use to elicit evidence of student learning. In addition to district and state assessments, they use learning checks like *Thumbs Up Thumbs Down* and consensograms, quizzes, questioning, debates, running records, peer grading, teacher-developed assessments and other classroom assessments.

Proficient CCI teachers also report improvements in the quality of their questioning and discussion strategies due to CCI. At the same time, some acknowledge this as an area ripe for further professional development. Research shows that effective questioning requires careful design so that teachers can identify undergeneralizations and overgeneralizations that students are known to make.¹² Effective techniques also involve adequate wait times

after questioning, and discussion strategies to prompt student thinking and gauge the understanding of the whole class rather than just selected students.¹³

Tests developed for summative and interim benchmark purposes may be used formatively, if time is set aside for analyzing results and then re-teaching and retaking tests.¹⁴ CCI provides the PDSA framework for using summative and benchmark tests in a formative way. Teachers and students graph test results from the A2L, KDPR, and DRA tests, and then identify strategies to increase proficiency levels.

Feedback that Moves Learning Forward

Research shows that much of the feedback that students receive has no impact on learning and can actually be counterproductive. A review of 131 scientifically rigorous studies suggests that “ego-involving” feedback, such as giving grades or scores, can reduce performance. Feedback appears most effective when it focuses on the quality of work and tells participants not just what to improve but *how* to improve it.¹⁵

Data collected from proficient CCI teachers in spring 2008 suggest that CCI encourages them to enhance their feedback practices. Among the changes they cite are: providing feedback more regularly; providing more specific and descriptive feedback; using written and verbal comments instead of or in addition to grades and test scores; and using more structured and more participatory feedback processes.

Commonly, however, CCI learning targets, assessment measures and feedback processes involve test scores and grades. CCI training materials provide this direction, and most APS schools report scores and grades. RQC may wish to examine this aspect of the CCI model to see if and how research-based feedback strategies could be more fully incorporated and diffused.

Activate Student Ownership

CCI asks students to take an active part in monitoring and regulating their own learning. Students monitor classroom and individual learning by graphing the results of tests, grades, quizzes, performance assessments, as well as learning activities, such as homework completion, attendance, and book reading. Using the PDSA process, students interpret assessment results and plan increasingly better ways to meet classroom and individual learning goals. Studies show that activating student ownership in this way can double the rate of learning, so that students learn in six months what students in control groups take a year to learn.¹⁶

Enhancements may include conducting self-assessments before, during and after instruction. For example, the *Traffic Lights* technique has students flash green, yellow or red cards to indicate their level of understanding. Such self-assessment methods not only provide teachers with important information, they also build students’ self-regulation skills.

Activate Peer Learning Resources

Research shows that activating students as learning resources for one another produces “some of the largest gains seen in any educational interventions,” provided that two conditions are met: (1) group goals, and (2) individual accountability.¹⁷ According to formative assessment research, applying these same conditions to collaborative assessment practices can improve student learning.¹⁸

CCI employs classroom, group and individual level PDSA processes to build collective responsibility for problem-solving and goal attainment. On a weekly basis, students strategize ways to meet group and individual goals. This often entails planning ways students can help each other learn, including for example, buddy strategies, peer tutoring, and practical supports, like calling classmates in the morning to ensure they get to school on time. RQC staff report that students hold each other accountable often more directly than would teachers.

Proficient CCI teachers report that CCI has expanded their use of peer learning strategies. Prior to CCI, some teachers used collaborative learning groups, peer tutoring, and small discussion groups. With CCI, more teachers report using peer strategies, the range of peer strategies is wider, and they use them more regularly.

An enhanced level of formative assessment practice would incorporate more peer assessment techniques, both verbal and written. Using the *Two Stars and a Wish* technique, for example, students provide two positive remarks and one suggested improvement. Peer assessment is uniquely valuable because students may accept critiques of their work from one another that they would not accept from a teacher, students have the chance to see things in others’ work that can enhance their own work, feedback is in a language that students themselves naturally use, and students learn by playing the roles of teacher and examiner.¹⁹

Innovation Adoption and Diffusion

The literature on innovation adoption and diffusion is another source of important insights about the CCI program. Like most innovations, CCI involves making significant changes in existing structures and practices. It therefore poses risks for teachers, administrators, students, parents, and others. Innovation adoption and diffusion models outline the CCI program’s likely implementation stages, challenges and timeline.

Concerns-Based Adoption Model

The Concerns-Based Adoption Model holds that innovation adopters progress through a series of adoption and implementation concerns, which must be addressed in sequential order. Before teachers are ready to receive training on how to implement an innovation, their awareness, informational and personal concerns must be addressed and resolved. Teachers then may need more than a year to resolve task-oriented concerns, such as how to use materials and tools and how to use time effectively. This phase involves experimenting with new approaches on a small scale, working out kinks, developing

confidence and then extending new approaches to other areas of practice. Only later can teachers focus on an innovation's impact on students, begin collaborating with other teachers in implementing the innovation and, finally, consider refinements that would enhance the innovation. The whole process of innovation adoption can take a minimum of three years for each set of adopters.²⁰

The Concerns-Based Adoption model substantiates the experiences of RQC's CCI program. Some CCI demonstration teachers resolved their informational and personal concerns very quickly and, within the program's first year, were able to focus on CCI implementation. Some of these teachers then resolved task-oriented concerns quickly. During the second program year they progressed to the next phase of extending CCI to multiple content areas, collaborating with other teachers and considering refinements.

Many teachers progressed more slowly. Some are not yet proficient, according to CCI proficiency assessments. They may harbor discouraging beliefs about CCI, their abilities or their school environments. These teachers, according to the CBAM, may not be ready or able to learn CCI processes and mechanics. Indeed, RQC provides proficient and non-proficient teachers with different training and support activities.

RQC may want to consider incorporating CBAM-related items into its professional development evaluation forms. One form comes from the PED and asks participants to rate sessions based on perceived usefulness, relevance and quality. New items could be added to assess participants' stage of concern so that RQC knows what issues need to be addressed for CCI adoption to occur. The CBAM framework also could help RQC chart participants' progress toward full CCI adoption.

Diffusion of Innovation

Diffusion of Innovation theory offers additional insights into CCI adoption and diffusion. RQC expects that each year the numbers of teachers implementing CCI proficiently will increase. Current implementers should continue and new implementers should begin.

Diffusion theory suggests that innovation adoption depends greatly on potential adopters' perceptions about the following innovation attributes:

- *Low Complexity*: Not difficult to understand and use;
- *Relative Advantage*: Better than status quo or alternative options;
- *Compatibility*: Consistent with current values and practices;
- *Observability*: Results and benefits are visible to others; and
- *Trialability*: Can be tried out on a partial or temporary basis.

Advocacy and training that address these attributes can facilitate innovation adoption. For example, RQC could consider ways to expand the observability of CCI practices and results. Discussions with school administrators and teachers could emphasize CCI's relative advantages, compatibility with existing practices, and ability to be tried on a limited basis.

Diffusion theory also suggests that CCI demonstration teachers and instructional coaches should be selected very carefully. Early CCI participants should be predisposed to innovation and should include respected opinion-leaders in each school. These 'Innovators' and 'Early Adopters' together represent about 16% of the general population. This translates to four innovators and early adopters in a school of 25 teachers. Their risk-tolerance is essential for modeling the innovation and producing visible results. Their credibility and leadership are essential to promoting adoption by the larger majority of more risk-averse and change-resistant individuals.

Unfortunately, RQC's attempts to recruit experienced and motivated teachers were thwarted from the beginning. Some teachers were sent to CCI trainings against their wishes and some were first-year teachers. None of these teachers could perform the work of Innovators, Early Adopters and opinion-leaders in their schools. This is a critical challenge for RQC to resolve in order to ensure CCI's adoption, diffusion and sustainability.

CCI Program Implementation

Evaluation Question #3: What do implementation experiences to-date tell us about the strengths and weaknesses of the CCI program?

“Desirable outcomes are achieved only when effective programs are implemented well.” The intervention’s core components must be clearly defined; staffing must be adequate, appropriate and well-supported; practitioners must use core intervention components and avoid contrary practices; and they must have enough skill to deliver core intervention components competently.²¹

In its second year of implementation, RQC’s CCI program is still in the stage of initial implementation. Most schools (26 out of 38) have deployment plans, however 12 do not. RQC reports that about half of the schools with plans (12) have not yet implemented their deployment plans successfully. Many teachers (66) and some instructional coaches (8) were rated CCI-proficient in 2007-08, but almost all are still learning and refining their CCI practice. Fourteen of the demonstration teachers and instructional coaches are not yet CCI-proficient. In addition to data collected by RQC, classroom visits provided RDA with direct evidence that CCI implementation varies widely across teachers, even within schools that have implemented CCI school-wide for multiple years.

Regular fidelity assessments and continuous improvement efforts are critical for achieving high levels of implementation.²² RQC’s fidelity systems include school-based deployment plans, site visits, practitioner meetings, self-assessment tools, a principal questionnaire, and a teacher questionnaire. These tools focus on measuring teachers’ compliance (with CCI processes and tools), competence (CCI delivery skills) and satisfaction.

Data collected from teachers by RQC staff suggest that teachers have varying levels of self-confidence in their ability to implement CCI, as well as varying levels of commitment to CCI. RQC staff report that teachers who implement CCI proficiently become convinced of its efficacy. They say the uncertain commitment of some teachers and ICs stems from poor implementation. Equally possible is that poor implementation stems from the uncertain commitment of some teachers, instructional coaches and school leaders. As referenced earlier, teacher assessment and professional development evaluation tools could be modified to gauge the scope and magnitude of these adoption concerns.

Contextual implementation factors are less well defined and assessed than compliance, competence and satisfaction. For example, schools set their own targets for how many teachers will be trained, and may or may not have adequate processes and resources for CCI supervision, evaluation and coaching. RQC offers school leaders support in developing deployment plans but, without clear targets, schools are likely to vary greatly in their plans and in their actual deployment of CCI.

In fact, RQC staffs report varying levels of commitment to CCI on the part of school leaders, a perspective supported by teacher questionnaire data collected by RQC. Some school leaders appear to see CCI as a key strategy to achieve AYP and improve the

school's accountability status. Other leaders may have agreed to implement CCI only as a concession to the PED.

RQC will need to continue its fidelity assessment and support processes in order to achieve a consistently high level of implementation. Only then can the program hope to produce measurable student achievement outcomes.

RECOMMENDATIONS

Recommendations for Program Enhancement

Evaluation Question #4: What programmatic enhancements, if any, could increase the CCI program's chances of achieving intended outcomes?

Provide teachers with more professional development in formative assessment strategies and techniques. The formative assessment literature is a rich source of information about how to use CCI to improve student achievement. The present study identifies some key ways RQC could reinforce and expand teachers' formative assessment practice, often within the context of CCI training and support activities. For example, CCI training and support could more specifically address both why and how to:

- Clarify students' understanding of the *quality* of work they are expected to produce;
- Use interim and benchmark assessments in a formative way;
- Use written and oral comments and questions as feedback to help students improve;
- Use embedded assessment and self-assessment techniques to guide instruction and build students' self-regulation skills; and
- Incorporate more peer assessment techniques, both verbal and written, into everyday classroom practice.

Leverage formative assessment research to promote CCI to district and school leaders. The district's explicit endorsement of CCI as a key district strategy is critical to CCI's ultimate success. RQC could cite CCI's links to formative assessment and the research showing increased student achievement at less cost than other popular strategies.

Incorporate Concerns-Based Adoption Model related items into teacher assessment and professional development evaluation tools in order to gauge training needs more accurately, and to be able to categorize participants into adoption stages.

Consider how Diffusion of Innovation theory can guide CCI adoption and diffusion efforts. For example, how can CCI practices and results be more visible? How can people be made more aware of CCI's relative advantages, compatibility with existing practices, and ability to be tried on a limited basis?

Develop a plan for recruiting teachers who possess Innovator, Early Adopter and Opinion Leader characteristics. These risk-tolerant and trend-setting individuals are key to diffusing CCI.

Focus CCI program and evaluation efforts on implementation rather than on outcomes for at least three years, the minimum time required for full implementation of

an innovation by any group of people. This should include establishing implementation monitoring spreadsheets or databases so that RQC staff can easily identify and communicate implementation gaps. It also means finding ways to maximize schools' compliance with implementation criteria, like how participating teachers are selected.

Clarify school-level implementation targets and establish school-level fidelity assessment processes. Classroom implementation of CCI depends in part on contextual factors, including supervision, coaching, time, and the availability of supportive resources. Specifying targets and adding them to RQC's existing fidelity assessment process would provide a more comprehensive picture of implementation and what RQC can do to promote fidelity.

Recommendations for Outcome Evaluation

Evaluation Question #5: What design elements, methods, resources and timeline are recommended for evaluating CCI program outcomes?

The goal of the CCI program is to improve reading and/or math student achievement in RQC demonstration schools and/or classrooms. The success measures proposed by RQC's funders are the percentages of schools or classrooms that show significant growth on the NM Standards Based Assessment and interim benchmark (Assess 2 Learn) tests. However, evaluating learning outcomes is much more complicated, as outlined below.

If RQC, JSA, NMPED and other stakeholders wish to evaluate the effects of CCI in a credible and meaningful way, a formal evaluation should be considered. Funding to support an evaluation would be necessary and may be available through stakeholder collaboration. The following pages recommend evaluation questions and suggest ways that RQC can prepare for such an evaluation.

Ultimate Outcomes

The key outcome evaluation question should address whether students who receive proficient CCI instruction are more likely to improve academically compared to similar students who do not receive CCI instruction. Assembling treatment and comparison groups must take account of both program dosage and fidelity of implementation. Students who have received *proficient* CCI instruction for a significant portion of the school year, and for significant portions of each school day, would be most likely to manifest measurable effects. Comparison students could be identified based on non-exposure to CCI and similarity to treatment students on factors such as prior academic achievement, grade level, free and reduced price lunch status, English language learning status, and special education status.

Another challenge is to collect sufficient implementation data to consider the wide range of internal and contextual factors that may influence the CCI program's ultimate goal. RQC collects some of these data already. Additional data that would be required to evaluate CCI outcomes include:

- Names and ID numbers for all students in CCI demonstration classes
- Teacher ID for each student, by semester or year
- Teacher's CCI proficiency level for each semester or year of participation
- Content areas in which CCI was applied, by semester or year
- Minutes per day of CCI application by quarter or semester
- Student involvement in remediation or tutoring programs
- Student attendance
- Teacher teaming and/or looping
- Proportion of teachers practicing CCI in each school
- CCI-related strategies that may have been implemented in the rest of the school, how long they have been implemented, and the number of teachers involved.

Table 2 shows a number of other outcome evaluation questions. In addition to measuring student learning, an evaluation also should explore the conditions under which positive outcomes are achieved. What amount of time, frequency of exposure and/or duration of exposure are associated with positive results? What aspects of CCI and/or ways of implementing CCI are most commonly associated with positive results? Answering these questions would require qualitative and quantitative data on school-level program implementation.

The formative assessment literature suggests that CCI may have cost advantages compared to other initiatives aimed at increasing student achievement. This possibility is worth investigating, although collecting necessary data would pose unknown challenges. What does CCI cost to produce increases in student achievement, in terms of staff time, training costs, materials and other budgetary items? Is there evidence that spending more produces better results? Are there other initiatives that cost more but achieve the same or less?

In addition, an evaluation should investigate the CCI program’s unintended and/or indirect consequences. For example, the CCI program may improve teachers’ job satisfaction and retention. If implemented school-wide, CCI may alter the school’s overall climate. It may also decrease student mobility, if parents feel more informed and excited about their children’s educational experience. On the other hand, CCI could be a bad fit for certain teachers and their effectiveness could decline as a result. Such possible unplanned consequences need to be factored into the overall evaluation of effects. Investigating unintended outcomes entails some open-ended qualitative inquiry to identify perceived benefits and detractors, followed potentially by quantitative methods to estimate their value or cost.

Table 2. Proposed Ultimate Outcome Questions and Methods

Ultimate Outcomes	Instruments
6.0. Are CCI students more likely than non-CCI students to achieve proficiency in reading and math?	SBA (Rdg & Math) Assess 2 Learn (Rdg & Math)
Under what conditions does the CCI model yield positive student results?	SBA, A2L, Administrator survey & teacher survey
What is the cost per student and how does that compare to other APS initiatives aimed at raising student achievement?	Expense records
What are the unintended consequences of CCI?	Teacher survey Administrator survey School climate surveys IC focus group Student focus group?

Intermediate Outcomes

To help explain ultimate student performance outcomes, an evaluation would need to address more proximal knowledge, behavior, learning and instructional outcomes. If students do not meet weekly learning goals, for example, they would not be expected to make significant gains on state and district tests. The evaluation would then need to explore reasons why students do not meet weekly goals. Data collection could include artifacts of student work, data folders, formative assessment artifacts and results, lesson plans, and a teacher survey. Proposed questions, their relation to the CCI Program Logic Model (Figure 1), and possible instruments are outlined in Table 3.

Table 3. Proposed Intermediate Outcome Evaluation Questions and Methods

Knowledge, Behavior & Instructional Outcomes	Instruments
5.0. What proportion of CCI students meets weekly learning goals?	Formative assessment data Data folders
4.1. What proportion of CCI students understands learning goals?	CCI quality tools
4.2. What proportion of CCI students takes responsibility for meeting learning goals (completes assigned and additional tasks)?	Artifacts and records of student work Data folders
4.3. Are teachers more knowledgeable about the effectiveness of their instruction and who has what learning deficiencies?	Teacher survey &/or focus groups Formative assessment artifacts
4.4. To what degree do teachers engineer extra learning activities to address students' learning deficiencies?	Lesson plans Artifacts & records of student work

Short-Term Outcomes

The degree to which students take responsibility for learning, and teachers' ability to identify and address learning deficiencies, depend on specific school and classroom practices. They also depend greatly on teachers' knowledge, self-efficacy and commitment to implement CCI. These short-term outcomes should be evaluated with a mix of quantitative and qualitative tools, including fidelity checks, site visits, document reviews and a teacher survey, as proposed in Table 4.

Table 4. Proposed Short-Term Outcome Evaluation Questions and Methods

School & Classroom Outcomes	Instruments
3.1 – 3. 5. At what level of quality and consistency, and in how many content areas, do CCI/PDSA processes occur?	CCI fidelity monitoring (site visits, self-assessments)
2.1. How many teachers total and per school have knowledge, self-efficacy and commitment to implement CCI?	CCI fidelity monitoring (site visits, self-assessments) Teacher survey

2.2. In how many classrooms is CCI practiced per school?	RQC site visit logs Principal report
2.3. How many schools have comprehensive, detailed CCI deployment plans?	Content analysis with quality rating matrix
2.4. What proportion of schools has leaders who understand CCI and support it with policies, resources, professional development and supervision/evaluation?	Teacher survey

Implementation Evaluation

In order to accurately attribute outcomes, or lack of measurable outcomes, implementation must be evaluated. This should include what activities were accomplished, by and for whom, when, and at what level of quality and dose. This same information is needed to identify the conditions under which the most positive outcomes seem to occur, as well as to identify possible improvements so more positive results can be achieved. A central implementation evaluation question would be: Under what conditions (school, classroom, teacher) does CCI adoption and implementation occur most efficiently and completely? Data collection could include activity logs or other existing records as well as a teacher survey, as shown in Table 5.

Table 5. Proposed Implementation Evaluation Questions and Methods

Implementation Activities	Instruments
1.1. What training, coaching, rewards and recognitions are provided to CCI demonstration teachers and ICs?	RQC records
1.2. What training, coaching and support activities do CCI demonstration teachers & ICs provide in their schools, how much and to whom?	Teacher survey Demonstration teacher/IC activity logs
1.3. What activities, with what participants, occur to facilitate the development of comprehensive, detailed school CCI deployment plans?	RQC records
1.4. What activities occur, with what participants, to educate school leaders about CCI & encourage supportive policies, resource allocations, professional development and supervision/evaluation?	RQC records

Assumptions

RQC's theory of change includes a number of assumptions that the evaluation must address in order to correctly attribute outcomes or explain a lack of outcomes. These include:

- Learning goals set by teachers are appropriate – promise to move students to or beyond proficiency in grade-level standards by the end of the school year.

- Learning activities are appropriate – engage students and match students’ learning needs and learning styles.
- Students attend school regularly.

Evaluation Timeline

The most cost-effective timing for outcome evaluation involves, first, establishing that a program is being implemented with a high level of fidelity and consistency. As discussed earlier, three or more years may be required to gain acceptance of an innovation, fully train staff, and establish consistent and quality practices. Additional time may be needed for the innovation, once established, to yield measurable effects. If attempts are made to measure outcomes prematurely, reviewers may conclude wrongly that a program has no effect.

Findings from this evaluation study suggest that RQC’s CCI program is in an initial implementation, rather than full operation, stage. In the short-term, therefore, RDA recommends that any evaluation should focus on short, intermediate, and implementation outcomes (Tables 3, 4 and 5). Additionally, RQC should continue its implementation monitoring and support activities, with modifications suggested in the Recommendations section of this report (pages 15-16).

Appendix A: CCI Proficient Teacher Self-Reports of Formative Assessment Practices Before CCI and With CCI, April 2008		
Formative Assessment Strategies	Pre-CCI Teacher Practices (n = 33 respondents)	CCI Practices
Clarify learning goals & criteria for success	<ul style="list-style-type: none"> Few communicated learning goals to students. None provided evidence of activities to ensure student understanding of goals. 	<ul style="list-style-type: none"> Most explicitly define learning goals, usually weekly. Many involve students in goal setting and performance measurement as ways of cultivating student understanding.
	<ul style="list-style-type: none"> Few used rubrics, examples of student work or defined measures to specify success criteria. 	<ul style="list-style-type: none"> Most use the PDSA goal format, which includes performance targets and measures. Some cite rubrics and curriculum mapping.
Conduct frequent assessments	Assessment cycles based on units, quarters or trimesters.	Most use a weekly assessment cycle for at least one content area.
Engage effective classroom discussions, questions, activities, and tasks that elicit evidence of student learning	<ul style="list-style-type: none"> Most used tests & quizzes but often for grades rather than to guide instruction or give students meaningful feedback. Some used discussions and questions, often provided by packaged curricula. Many cited observation. 	<ul style="list-style-type: none"> Teachers report a wider array of tools and techniques for measuring learning. Teachers report more effective questioning and discussion strategies.
Provide feedback that moves learning forward	<ul style="list-style-type: none"> Most commonly grades and some comments are given as feedback. Verbal feedback may occur only at student & parent conferences. 	<ul style="list-style-type: none"> Feedback occurs more regularly. Feedback provides more specific guidance. Teachers provide more written and verbal comments Teachers use more structured and more participatory feedback processes
Activate students to take ownership of their own learning	May ask students to take responsibility for their own learning but no structures or processes are provided.	<ul style="list-style-type: none"> Students use CCI tools to monitor their own progress and class progress. Using PDSA process, students plan increasingly better ways to meet learning goals.
Activate students as learning resources for one another	Some use of collaborative learning groups, peer tutoring, and small discussion groups	Wider range of peer learning strategies, used more regularly and by more teachers.

End-Notes

- ¹ Arter, J. & Stiggins, R. (2005, September). National Council on Measurement in Education Newsletter, 13(3), 4.
- ² Black, P., Wiliam, D. (1998). Inside the Black Box: Raising Standards Through Classroom Assessment. *Phi Delta Kappan*, 80 (2). Retrieved April 22, 2008 from <http://www.pdkintl.org/kappan/kbla9810.htm>.
- ³ Wiliam, D. (2006, July) *Does Assessment Hinder Learning?*. Speech delivered at the ETS Breakfast Salon. Retrieved April 21, 2008 from http://www.uk.etseurope.org/fileadmin/free_resources/UK%20website/UK_Current/Dylan_Wiliam_speech.doc
- Yeh, S. (2007). The Cost-Effectiveness of Five Policies for Improving Student Achievement. *American Journal of Evaluation*, 28(4), 416-436.
- ⁴ Shepard 2007, as cited in Wiliam, D. (2007). *What Does Research Say the Benefits of Formative Assessment Are?* Assessment Research Brief National Council of Teachers of Mathematics.
- ⁵ Chappuis, J. (2005). Helping Students Understand Assessment. *Educational Leadership*, 63(3), 39-43. Retrieved April 22, 2008 from <http://www.iowa.gov/educate/content/view/1124/1490>
- ⁶ Wiliam, D. (2007). *Five Key Strategies for Effective Formative Assessment*. Assessment Research Brief. National Council of Teachers of Mathematics.
- ⁷ Sadler 1989, as cited in Wiliam, D. (2007). *Five Key Strategies for Effective Formative Assessment*. Assessment Research Brief National Council of Teachers of Mathematics.
- ⁸ White and Fredericksen 1998, as cited in Wiliam, D. (2007). *Five Key Strategies for Effective Formative Assessment*. Assessment Research Brief. National Council of Teachers of Mathematics.
- ⁹ Wiliam and Thompson, 2007, as cited in Wiliam, D (2007). *What Does Research Say the Benefits of Formative Assessment Are?* Assessment Research Brief National Council of Teachers of Mathematics.
- ¹⁰ Popham, W. J. (2006). Phony Formative Assessments: Buyer Beware! *Educational Leadership*, 64(3), 86-87.
- ¹¹ Stiggins, R. (2007). Assessment through the student's eye. *Educational Leadership*. 64(8), 22-26. Retrieved April 22, 2008 from <http://www.iowa.gov/educate/content/view/1124/1490>
- ¹² Bransford, Brown & Cocking, 2000, as cited in Wiliam, D. (2007). *Five Key Strategies for Effective Formative Assessment*. Assessment Research Brief. National Council of Teachers of Mathematics.
- ¹³ Wiliam, D. (2008). Changing Classroom Practice. *Educational Leadership* 65(4), 37.
- ¹⁴ Chappuis, S. & Chappuis, J. (2008). The Best Value in Formative Assessment. *Educational Leadership* 65(4), 16.

¹⁵ Kluger & DeNisi, 1996, as cited in Wiliam, D. (2007), *Five Key Strategies for Effective Formative Assessment*. Assessment Research Brief. National Council of Teachers of Mathematics.

¹⁶ Fontana & Fernandes 1994; Mevarech & Kramarski 1997, as cited in Wiliam, D. (2007). *Five Key Strategies for Effective Formative Assessment*. Assessment Research Brief National Council of Teachers of Mathematics.

¹⁷ Slavin, Hurley & Chamberlain, 2003, as cited in Wiliam, D. (2007). *Five Key Strategies for Effective Formative Assessment*. Assessment Research Brief. National Council of Teachers of Mathematics.

¹⁸ Black,P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*, 86 (1), 14.

¹⁹ Black,P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*, 86 (1), 14-15.

²⁰ Loucks-Horsley, S. (1996). *The Concerns-Based Adoption Model (CBAM): A Model for Change in Individuals*. Reprinted from the chapter entitled, Professional Development for Science Education. In *National Standards and the Science Curriculum*. Dubuque, Iowa: Kendall/Hunt Publishing Co. Retrieved April 24, 2008 from <http://www.nas.edu/rise/backg4a.htm>

²¹ Fixsen, D.L., Naoom, S.F., Blase, K.A., Friedman, R.M., Wallace, F. (2005). *Implementation Research: A Synthesis of the Literature* (FMHI Publication #231). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network, p. 12.

²² Fixsen, D.L., Naoom, S.F., Blase, K.A., Friedman, R.M., Wallace, F. (2005). *Implementation Research: A Synthesis of the Literature* (FMHI Publication #231). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.