Teacher Self-Assessment: A Mechanism for Facilitating Professional Growth

John A. Ross*

Ontario Institute for Studies in Education

University of Toronto

Cathy Bruce

Trent University

Presented at AERA, Montreal

April, 2005

*Corresponding Author:
Dr. John A. Ross,
Professor of Curriculum, Teaching, & Learning,
OISE/UT Trent Valley Centre,
P.O. Box 719, 1994 Fisher Dr.,
Peterborough, ON K9J 7A1 Canada
Tel: 705-742-9773, ext 2293; Fax: 705-742-5104; jross@oise.utoronto.ca

Teacher Self-Assessment: A Mechanism for Facilitating Professional Growth

Abstract

Self-assessment is a powerful technique for improving achievement. In this article we outline a theory of teacher change that links self-assessment by teachers to their professional growth. This theory provides avenues for peers and change agents to influence teacher practice. We apply the theory to change in mathematics teaching and report an explanatory case study in which use of the self-assessment tool, in combination with other elements, contributed to change in the instructional practice of a grade 8 mathematics teacher. Provision of a self-assessment tool contributed to teacher growth by: 1) influencing the teacher's definition of excellence in teaching and increasing his ability to recognize mastery experiences; 2) helping the teacher select improvement goals by providing him with clear standards of teaching, opportunities to find gaps between desired and actual practices, and a menu of options for action; 3) facilitating communication with the teacher's peer, and 4) increasing the influence of external change agents on teacher practice. The study argues that providing a self-assessment tool is a constructive strategy for improving the effectiveness of in-service provided it is bundled with other professional growth strategies: peer coaching, observation by external change agents, and focused input on teaching strategies.

Teacher Self-Assessment: A Mechanism for Facilitating Professional Growth[1]

Self-assessment is a powerful technique for self-improvement. Teaching students how to self-assess contributes to more accurate self-assessment and to higher student achievement (Arter, Spandel, Culham, & Pollard, 1994; McDonald & Boud, 2003; Ross, Rolheiser, & Hogaboam-Gray, 1999; 2002-b; Ross & Starling, in press). In this article we outline a theory of teacher change in which self-assessment is a mechanism for professional growth that provides avenues for peers and change agents to influence teacher practice. Because it provides opportunities for influence, it is a hitherto neglected strategy for facilitating professional growth. We apply the theory to change in mathematics teaching and report an explanatory case study in which use of the self-assessment tool, in combination with other elements, contributed to change in a grade 8 teacher.

Theoretical Framework

Theory of Individual Teacher Change

We began the study with the model of teacher change in Figure 1, a model heavily influenced by social cognition theory (Bandura, 1997) and by our work on student self-assessment (particularly Ross et al., 2002-b). It proposes that teacher change occurs through reflection on experience and that self-efficacy beliefs mediate the influence of self-assessment on teacher practice.

Figure 1 Here

[1] The research was funded by the Ontario Ministry of Education and Training and by the Social Sciences and Humanities Research Council. The article does not necessarily reflect the views of the Ministry or the Council. The rubric for standards-based mathematics teaching and the self-assessment tool were created by John Ross, Doug McDougall, Anne Hogaboam-Gray, Cathy Bruce, Sonia Ben Jafaar, Jane Lee and Ann LeSage, with website design by Carolyn Brioux.

At the centre of Figure 1 are teacher cognitions about their practice, particularly their perceptions of its effectiveness in bringing about student learning. The model represents self-assessment as the integration of three processes that self-regulating professionals use to observe and interpret their behavior. First, teachers produce self-observations, focusing on aspects of instruction relevant to their subjective standards of success. Second, teachers make self-judgments in which they determine how well their general and specific goals were met. The primary data are teacher perceptions of changes in student performance gleaned from student utterances, work on classroom assignments, homework, and formal assessments. Third, are self-reactions, interpretations of the degree of goal attainment that express how satisfied teachers are with the result of their actions. These self-assessments may occur in the moment (if the duration of the moment is extended) as reflection-in-action or retrospectively as reflection-on-action.

Self-assessments contribute to teachers' beliefs about their ability to bring about student learning; i.e., teacher efficacy, a form of professional self-efficacy. Teacher efficacy is particularized to teaching specific content, to particular students, in specific instructional contexts. Teacher efficacy is an expectancy about future performance that is based on past experience. Teachers who perceive they have been successful, regardless of the accuracy of their judgment, expect to be successful in the future. Of the four sources of teacher efficacy information identified by Bandura (1997), the most powerful is mastery experience—being successful in the classroom. Teachers become confident about their future performance when they believe that through their own actions they have helped children learn.

Teacher efficacy influences goal setting and effort expenditure. Teachers who anticipate that they will be successful set higher goals for themselves and their students. Teacher efficacy consistently predicts willingness to try out new teaching ideas, particularly techniques that are

difficult to implement and involve risks such as sharing control with students (e.g., Ross, 1992-b). High expectations of success motivate classroom experimentation because teachers anticipate they will be able to achieve the benefits of innovation and overcome obstacles that might arise. Teachers with high expectations about their ability to teach produce higher student achievement (Goddard, Hoy, & Hoy, 2004; Herman, Meece, & McCombs, 2000; Mascall, 2003; Muijs & Reynolds, 2001; Ross, 1992-b; Ross & Cousins, 1993). Teacher efficacy contributes to achievement because high efficacy teachers try harder, use management strategies that stimulate student autonomy, attend more closely to low ability student needs, and modify students' ability perceptions (evidence reviewed in Ross, 1998).

Teacher efficacy also influences effort. Teachers with high teacher efficacy have fewer absences (Imants & Van Zoelen, 1995), are more willing to handle difficult to teach students themselves rather than refer them to special classes (Soodak & Podell, 1998), and are less likely to leave the profession (Brouwers & Tomic, 2000; Johnson & Birkeland, 2003). Confident teachers persist. They are not depressed by failure but respond to setbacks with renewed effort (Bandura, 1997).

Goal setting and effort expenditure are linked: For example, individuals are more likely to persist if they adopt goals that have unambiguous outcomes, that are achievable in the near future, and that are moderately difficult to achieve (Schunk, 1981).

The combination of goals and effort affect teacher practice defined as the assembly of teacher actions (e.g., choice of curriculum objectives, teaching methods, assessment) and knowledge (e.g., of subjects, learners, pedagogy, and policy). Teachers willing to try new instructional ideas and persist through obstacles are more likely to implement new approaches, experience success with them, and internalize the innovations into their practice. There may also

be benefits for professional learning. Research with children suggests that positive self-evaluations foster an upward cycle of learning in which the child sets higher goals and commits more personal resources to learning tasks. Negative self-evaluations lead students to adopt social rather than academic self-images (Anderman, Anderman, & Griesinger, 1999), embrace goal orientations that conflict with learning, select personal goals that are unrealistic, adopt learning strategies which are ineffective, exert low effort, and make excuses for performance (Stipek, Recchia, & McClintic, 1992).

In summary, our theory of teacher change is based on teacher self-assessment within the broader framework of social cognition theory. What matters is not the absolute level of classroom success but teachers' interpretation of experience. Self-assessment contributes to expectations that guide goal setting and effort.

The Influence of Peers on Teacher Self-Assessment

Self-assessment, an individual enterprise, can be informed by colleagues. Peer input can influence the first self-assessment process (self-observation) by directing teacher attention to particular dimensions of practice. Peer feedback might influence teacher judgments about the degree of their goal attainment (the second process). The influence is likely to be stronger if colleagues interpret student outcomes and teacher practices in terms of overt standards. Peers might influence teacher satisfaction with the outcomes of their instruction (the third process), if colleagues give praise explicitly linked to the quality of the teacher's performance (Cameron & Pierce, 1994). Peer feedback is weighed against the observations, judgments and reactions the teacher generates during (i.e., reflection-in-action) and after (i.e., reflection-on-action) the lesson. Peer input may complement or compete with these self-responses depending upon the teacher's perceptions of the credibility of their colleagues.

These opportunities for peer influences on teacher self-assessment that contribute to self-efficacy involve recognizing teaching success (mastery experiences). Peers also have opportunities to influence teacher efficacy through the other three sources of efficacy information proposed by Bandura (1997): social persuasion (telling colleagues they are capable of performing a task), vicarious experience (highlighting the successful performance of someone similar to the teacher), and managing physiological and emotional states (strengthening positive feelings arising from teaching and interpreting them as indicative of teaching ability or reducing negative feelings arising from teaching, such as stress).

Peers can also influence teacher practice by suggesting specific strategies and by working together to implement them. Collaboration among teachers promotes teacher efficacy (e.g., Chester & Beaudin, 1996), especially when it leads to instructional coordination within a school (e.g., Raudenbush, Rowan, & Cheong, 1992).

Norms of privatized practice typically limit peer opportunities for influencing teacher self-assessments. Norms of isolation can be overcome by creating professional school communities with shared values, collaborative action, deprivatized practice and reflective dialogue (Louis & Marks, 1998). A structured approach for doing this is peer coaching--pairs of teachers of equal experience and competence observe each other teach, negotiate improvement goals, devise strategies to implement the goals, observe the improved teaching, and provide each other with feedback. Coaching has positive effects when the appropriate climate, involving mutual trust, genuine voluntarism, encouragement of reflective thinking, and principal support (McLymont & da Costa, 1998) is developed. Peer coaching increases teacher implementation of sought-after teaching practices and contributes to higher teacher efficacy (Kohler, Ezell, & Paluselli, 1999; Licklider, 1995; Wineburg, 1995).

When teachers reinforce each other, beliefs about competence are magnified. Collective teacher efficacy is "the perceptions of teachers in a school that the efforts of the faculty as a whole will have a positive effect on students" (Goddard, Hoy, & Hoy, 2000, p. 480). Schools with high collective teacher efficacy have higher student achievement than schools with lower levels of collective teacher efficacy, independent of student socio-economic status (e.g., Goddard, 2001; 2002b; Goddard et al., 2000). Peers exert a powerful influence on the collective teacher efficacy of a school.

In summary, the individual process of teacher change can be influenced by teachers' peers. The avenues for influence are embedded in the self-assessment process and in other mechanisms identified in social cognition theory such as persuasion and vicarious experience. But these avenues need to be developed through professional community development activities such as peer coaching.

Influence of External Change Agents on Teacher Change

External change agents, for example, university researchers, can influence teacher self-assessment through the same mechanisms potentially available to peers. Although lack of immediacy, presence, and shared values diminishes the influence of external agents, university-based change agents can contribute to teacher self-assessment by clarifying goals (i.e., dimensions of teaching that define excellence) and criteria (levels of performance that constitute a hierarchy of professional growth). They can also provide credible feedback on whether particular standards of teaching have been achieved (e.g., through classroom observation). Researchers can also provide teachers with information about means—specific strategies that are credibly linked to subject standards that are likely to lead to higher achievement

Application to Mathematics

We selected mathematics as the domain for testing our theory of teacher self-assessment because the field distinguishes clearly between standards-based teaching and traditional methods; there is evidence that the former has greater student achievement effects than the latter; and implementation of math education reform is a challenging school improvement problem. With implementation of standards-based mathematics teaching, problem solving and conceptual understanding improve without any loss of computational mastery (Hickey, Moore, & Pellegrino, 2001; Riordan & Noyce, 2001; Wood & Sellers, 1997). Yet research (reviewed in Ross, McDougall, & Hogaboam-Gray, 2002, 2002) suggests that change in mathematics classrooms is problematic. For example, teachers must be agents of a change they did not experience as students (Anderson & Piazza, 1996).

Our review of math education research (Ross et al., 2002) and National Council of Teachers of Mathematics policy statements (NCTM 1989; 1991; 2000) identified ten characteristics of standards-based mathematics teaching. We constructed a rubric from observations and interviews with teachers who ranged from traditional to high fidelity implementers of standards-based mathematics teaching (McDougall, Lawson, Ross, MacLellan, Kajander & Scane, 2000; Ross, Hogaboam-Gray, McDougall, & Bruce, 2001; Ross & McDougall., 2003). For each of our 10 dimensions, we described four levels of implementation, arranged in a hierarchy of increasing fidelity to NCTM Standards. Table 1 illustrates the rubric. The Table shows four levels describing the "transparency" of student evaluation; i.e., the extent to which the teacher's procedures and criteria for evaluation are revealed to students. The validity of the hierarchy of levels was established by a panel of content experts (Ross et al., 2003). We created a self-report survey from the rubric to produce evidence of the validity of the

survey and the rubric. The evidence consisted of correlations of survey scores with a mandated performance assessment, congruence with classroom observations of teachers, and demonstrations that teachers who were similar in their claims about using a standards-based text series differed in how they used the text, in ways predicted by the survey (Ross et al., 2003).

The self-assessment tool consists of four short descriptions (derived from the rubric) for each of the 10 dimensions of Standards-based mathematics teaching. Teachers access the tool through an interactive website (URL removed for blind review). The teacher makes 18 decisions and receives immediate feedback. The website provides teachers with an overall score in terms of four categories ranging from "procedures focus", corresponding to traditional mathematics instruction, to "exploration focus", corresponding to standards-based mathematics teaching. Teachers also receive a summary of their responses to all items.

Case Study

We tested the utility of self-assessment as a device for teacher change with data collected in a qualitative study of ten grade 7-9 teachers engaged over a three-month period in an in-service program to increase implementation of standards-based teaching (Ross & McDougall, 2003). The intervention had four components: 1) teacher self-assessment and 2) the teaching rubric, both described above. 3) Information on how to implement the Standards based on the PD Standards for Elementary Mathematics identified by Hill (2004): teachers constructed mathematical meaning by engaging in mathematical tasks and content comparable to those undertaken by their students; the in-service focused on classroom practice (e.g., teachers examined examples of student work from their own classes and classes of their peers); teachers worked together, rather than alone, on in-service tasks; in-service presenters modeled the recommended instructional practices; the in-service focused on student learning, illustrating how

students learn mathematics; classroom teachers participated in the design and delivery of the in-service. These standards contribute to teacher learning (Brandes & Erickson, 1998; Garet, Porter, Desimone, Birman, & Yoon, 2001; Loucks-Horsley & Matsumoto, 1999). 4) Peer coaching in which pairs of teachers observed each other teach, collectively set professional goals, implemented the goals individually, and observed changes in each other's classroom practice. Coaching has positive effects when the appropriate climate, involving mutual trust, genuine voluntarism, encouragement of reflective thinking, and principal support (McLymont & da Costa, 1998) is developed. Peer coaching increases teacher implementation of sought-after teaching practices and contributes to higher teacher efficacy (Kohler, Ezell, & Paluselli, 1999; Licklider, 1995; Wineburg, 1995). The events of the intervention were: 1) individual self-assessment using an interactive website; 2) in-service on peer observation skills and teaching strategies for Standards application; 3) peer observation of teaching, 4) in-service on use of peer observation data and input on teaching strategies; 5) classroom experimentation over 4 weeks; 6) peer observation, and 7) in-service on teaching strategies.

For this article we purposefully selected the teacher who most benefited from the intervention. The study is an explanatory case study; i.e., it provides the best evidence that we have of the cause-effect relationships represented in Figure 1. The mechanisms that were operating in our explanatory case were present in the other cases, albeit to a lesser degree (see URL removed for blind review). We used pattern matching (Mark, Henry, & Julnes, 2000) in which we compared hypothesized to observed events to test the claim that self-assessment contributes to professional growth. In doing so we contrasted the mechanisms hypothesized in Figure 1 with a simpler model in which teacher change was attributed to providing teachers with information on how to teach mathematics from a standards-based perspective. This simpler

model assumes that teachers are not implementing reform mathematics because they lack pedagogical content knowledge.

Methods

Barry was a grade 8 teacher in a grade 7-8 Ontario, Canada school who taught five math classes and a homeroom class. The school served 300 urban and rural pupils from a range of economic backgrounds. Barry was a seasoned mathematics teacher, recognized for his expertise in the school and district. His teaching partner was Mark, a grade 7 teacher in the same school who taught several Grade 7 mathematics classes and a homeroom class.

We observed Barry on five occasions (observations 1 and 5 were two day events; the remainder were one day).Site visits followed procedures established by Simon and Tzur (1999). We visited Barry's classroom in September on two consecutive days during his math period (75-80 minutes per day). We interviewed him before, during, and after each math lesson to elicit his intentions and reflections on the lessons we observed. We recorded key events and contextual detail. The purposes of the second day of observations were: a) to ensure that as many dimensions of teaching as possible were observed; b) to determine Barry's consistency; c) to guard against demonstration lessons. Observations of mathematics teaching (by Barry, his peer, and researchers) were coded using the categories of the rubric (4 levels X 10 dimensions of teaching), using the template shown in the Appendix. The data consisted of the self-assessment, our classroom observations, Mark's peer observation reports, individual and collective interviews, correspondence with Barry, Barry's response to our 10-page case report, and field notes of in-service sessions.

Analysis was guided by three questions: In what ways did Barry think he had changed? To what extent did our observations confirm or disconfirm Barry's claims about change? What

factors contributed to or impeded change? We developed codes (N=28), shown in the Appendix, based on the questions and used NUD*IST to organize the data. Themes were developed through constant comparison. Credibility of the findings was enhanced by 1) triangulating between data collection times and interpreters (Creswell, 1998); 2) maintaining an audit trail by creating charts of relationships and counting instances (Miles & Hubberman, 1994); 3) searching for evidence of alternative theories; i.e., testing the alternate hypothesis that provision of pedagogical content knowledge is sufficient for teacher change (Mark et al., 2000); 4) using member checks, particularly teacher feedback on the short case reports (Creswell, 1998).

Results

Barry's Action Plan Barry's self-perception on entry to the project and the results of his web-based self-assessment were in agreement, leading Barry to summarize his teaching approach as: "The kids explore, but mostly on my terms. I'm at a level 3 (modified exploration focus) with some signs of level 4 (exploration focus)." Initial observations matched this interpretation. Barry encouraged problem solving and integration of different types of mathematics using rich learning tasks. For example, students applied measurement, geometry and number concepts and skills to design a camp. He deepened understanding through multiple and varied opportunities to explore specific concepts and skills. For example, after conducting surveys and participating in a whole group data management task, students were given a series of statements to explore, such as "On average, Canadians open their refrigerators 22 times a day." Students had to create a mathematical rationale as to whether they thought each statement was reasonable or unreasonable.

Barry chose to focus on Dimension 8: Student-Student Interaction and Dimension 9: Student Assessment (especially transparency of assessment, illustrated in Table 1), seeing these

as relative weaknesses in his program. This article will focus on the student assessment aspect of his action plan as an illustrative example.

During our initial observations Barry used a range of high quality assessment strategies. Each student kept a portfolio containing rubrics for assessing specific tasks, anecdotal comments made by the teacher on student work, end of unit student reflections and end of unit culminating tasks. A typical written comment included: "I'm glad you're enjoying math so far. You need a more complete write-up of the hummingbird experiment." We observed Barry using a rubric while students presented their data displays in class. During the initial interview, Barry indicated that he provided students with rubrics for performance assessments, but not for tests. He provided examples of his practice that indicated that his purposes for student assessment included reporting to parents, improvement of student learning, refinement of teaching methods and curriculum modification. Barry's self-assessment was confirmatory in that it placed him toward the high end on the subdimensions "variety of assessments" and "purpose of assessments" but it also highlighted an aspect of Barry's practice that he wanted to change. Although Barry used assessment rubrics he did not involve students in their development, something he saw Mark doing during the peer observation. Barry's plan was to develop assessment tools collaboratively with students, making assessment a more transparent process, and shifting from a teacher-directed assessment to more student-directed assessment. Mark expressed confidence that Barry would be successful in changing his assessment practice.

Implementation of Barry's Action Plan Barry embarked on a Fermi Problem Solving unit. (A Fermi problem generally requires estimation of physical quantities. These problems challenge students to ask more questions, define the parameters of their solution and determine reasonableness, not just provide an answer.)

Barry began the unit by negotiating with students a rubric for problem solving, adapting a strategy presented at the second in-service session. His strategy had three phases. In the first phase, students individually brainstormed responses to questions like "when I am solving problems in Math, it is helpful when…If our class was problem solving in small groups, an observer would hear…" In groups of 3-4 without teacher intervention, students created graffiti charts to organize and display their responses. In the second phase, students worked in groups to solve a rich problem. During the 75 minutes it took to complete the problem four students observed problem solving behaviors. The observers used a T-Chart labeled: "What I hear in the problem solving setting" and "What I see in the problem solving setting". In the third phase, Barry shared a set of five criteria for assessing student problem solving (from provincial assessment policy). In small groups, students created descriptors for each criterion, drawing upon the vocabulary of their graffiti and T-charts. The final rubric was posted across one full wall in the classroom. Barry told students that they would be evaluated based on these descriptions.

In subsequent classes we observed Barry opening the lesson by reviewing the criteria for evaluation of problem solving activity.

T: What will I be looking for today?

St: On task, quiet work.

T: What kind of work?

Sts: Various students referred to specific components of the rubric on the side wall to describe the kinds of work the teacher would be looking for.

T: So I will be looking to see how you work with your partner. I'll be doing that on my tracker.

After this class, the teacher wrote a note to each small group about his observations of their interactions and related the observations to the class rubric. For example:

> Jessica and Soroya [level 3 of the student rubric, dimension "solution write up"]
>
> You were both very focused during my observation and were sharing ideas and building on the ideas of others. 6 solutions was quite a good number for the time allotted.

Students were observed reviewing these teacher comments very carefully.

In another class we observed Barry telling students at the start-up which items from the student negotiated rubric he would focus on during his observations of group work. There was a one minute review of the rubric on the wall before students began their work. During this observation, student assessment was made explicit and was based on student negotiated criteria.

Teacher Perceptions of Professional Change When asked whether his practice had changed with regard to the transparency of student assessment, Barry said "Over the four weeks [of intervention and observations], I opened things up. In May, the survey would change again. For example, I think there will be more transparency." He identified evidence of the effects of his change on students, such as "Individual student results in problem solving increased from the previous unit, students were faster getting on task [times were recorded]", and student write-ups [written responses to problems] were more elaborate."

After the end of the project Barry continued to work on assessment transparency, as indicated by his notes to the researchers. In April, for example, he reported that student work was scored on a student negotiated rubric for a unit based on the theme of the human body. In the same month, Barry gave a workshop to teachers in his district on his strategies for creating transparency in student assessment. In June Barry reported: "Over the summer, I hope to refine some of the tasks that we used in the earlier phases of the project and use them right at the start

of September…so that the message to the students is clear that I value transparency and a sharing of learning on the mathematical road."

We agreed with Barry's judgment. Barry began the project working at a modified exploration focus level for the transparency sub-dimension of student assessment (level 3 in Table 1). After implementation of his action plan, Barry had shifted his program to an exploration focus on this dimension (level 4).

Discussion of the Case

Self-Assessment as an Enabling Factor in Professional Growth No previous study has explored teacher self-assessment as a mechanism for teacher change. The data from the case provides considerable support for the mechanisms in Figure 1. First, the web-based self-assessment affected Barry's internal self-assessment by modifying his definition of excellence in mathematics teaching. The self-assessment tool identified specific dimensions he could attend to in judging his work and provided descriptions of teaching to which he could compare his performance. The articulation of criteria for effective mathematics teaching enhanced Barry's beliefs about his competence. Barry scored toward the high end on almost all the scales, including the dimensions he chose to work on. (His self-assessment and researcher observations placed him slightly lower on the student-student interaction dimension but by the end of the study Barry had made visible improvements on it as well.) By modifying Barry's definition of teaching excellence the self-assessment increased Barry's ability to recognize his success. The credibility of the self-assessment was enhanced by its agreement with researcher and peer observations. In social cognition terms, the self-assessment tool contributed to two sources of teacher efficacy information: mastery experiences and persuasion.

Second, the self-assessment tool helped Barry select improvement goals by providing him with clear standards of teaching, opportunities to find gaps between desired and actual practices, and a menu of options for action. Barry believed that the rubric for standards-based teaching accurately represented the domain of mathematics education and he agreed with the hierarchy in which the descriptions of teaching were arranged. He was able to place himself within this scheme, recognize discrepancies between his current placement and where he would like to go, and select which gaps to work on first. He perceived the specific goals that he identified as immediate and feasible. As Barry implemented his action plan he experienced success. He saw how his actions led to student outcomes he valued. This strengthened Barry's beliefs about his competence and renewed his commitment to standards-based teaching.

Third, the self-assessment tool, particularly the rubric, facilitated communication between Barry and Mark. It provided them with a common language for talking about teaching (i.e., the dimensions), common scales (i.e., the levels of the rubric), and a shared experience (i.e., the novel experience of comparing their teaching to a variety of scales and receiving feedback on their assessments from peers and researchers). These mechanisms increased the clarity and credibility of the advice Mark offered to Barry. In addition Barry's observations of Mark provided Barry with assurance that he was *"*on the right track*"*.

Fourth, the self-assessment tool increased the influence of external change agents on Barry, through similar mechanisms identified for peer influence. The key influence was exercised by providing clarity about standards of teaching, which contributed to Barry's self-efficacy beliefs and to his goal setting. Barry's self-efficacy was further heightened by the researchers' observations and feedback on his practice. In addition the self-assessment data enabled the researchers to target dimensions of teaching that were most relevant to Barry's plan

and to calibrate the input so that it was in Barry's zone of proximal development (Vygotsky, 1978).

Self-assessment as a Constraint on Professional Growth Barry thought that the self-assessment tool was too rigid. Although each set of response options in the self-assessment is a continuum, the feedback given by the website implied that the underlying metric is an ordinal scale. Barry felt his practice overlapped the categories and this lowered the credibility of the self-assessment. In addition, Barry felt that the rubric levels represented such large changes in practice that the web-based tool was unsuitable to measure change over short time periods.

Refinement of Our Model of Teacher Change Data from the case led us to refine the model in three ways, as shown in Figure 2. First, Barry reported that the changes that he made in his instructional practice contributed to improved student achievement (e.g., higher scores in problem solving), independent of any changes that occurred in his standards for interpreting the quality of their work. We decided to include student achievement, measured independently of teacher interpretations, as an element in our model that is a consequence of instructional change and an antecedent of teacher self-assessment. Second, we found that both peers and researchers contributed to knowledge of innovative instruction. Researchers contributed to teachers' pedagogical content knowledge through semi-structured in-service sessions that demonstrated constructivist teaching strategies associated with mathematics education reform. Peers contributed teaching strategies during sharing sessions in which they described their classroom experiments. They also responded to researcher presentations with adaptations that peers had implemented, read about, or invented on the spot. The assembly of an array of instructional alternatives contributed to goal setting and to effort allocation in the sense that teachers could focus on adapting specific suggestions as opposed to creating new teaching strategies out of

whole cloth. Third, we recognized that the self-assessment needed to identify deficiencies in present practice (in Barry's case the discovery of lack of transparency in his assessment practice), otherwise there was no reason for teacher's to change.

Figure 2 About Here

Negative Data We selected Barry as our explanatory case because his experience best illustrated the causal relationships we wanted to explore. Our rationale was that if these relationships could not be demonstrated in optimal conditions they are probably too weak to pursue. Six of the other nine cases in the study also manifested positive, albeit smaller changes in practice (see Ross & McDougall, 2003). In all cases, we saw some evidence of the effects of the self-assessment tool. For example, all teachers, regardless of their commitment to the project, used the rubric for standards-based teaching to set goals and gauge change in their teaching.

Each of the three cases in which teachers did not change in the directions increased our understanding of how self-assessment influences teacher practice. One teacher's self-assessment was higher than was indicated by peer and the researcher observations. This teacher rejected attempts to provide him with contrary evidence, enjoyed the in-service and discussions about mathematics teaching, and emerged from the project virtually unchanged (in our view—he reported that he had). Our interpretation was that his beliefs about mathematics teaching were compatible with the assumptions of the project but he never saw the gap between his beliefs and practice and for that reason did not attempt to change. His instructional experiments were replications of existing practice. Two other teachers, working as a pair, interpreted their self-assessments positively by reversing the direction of the rubric levels—they saw the traditional focus pole of each dimension as more desirable than the exploration-focus pole represented by reform. In the peer coaching they consolidated their professional values and constructed a pool

of traditional instructional practices that facilitated their movement away from practices recommended in the provincial curriculum. These negative cases refined our model in that we came to see teacher change as the outcome of self-assessment grounded in credible data and defensible standards, dissatisfaction with present performance, access to powerful instructional alternatives, and confidence about one's ability to implement new instruction. In addition we noted that these conditions were associated with contextual elements that enabled Barry to benefit from the self-assessment tool: high congruence between his beliefs about mathematics teaching and the beliefs implicit in the self-assessment tool, a history of successful math teaching, a strong collegial relationship with a peer, and strong support from school administration.

Conclusion: Provision of a Self-Assessment Tool is Constructive But Insufficient

The explanatory case illustrates how self-assessment can contribute to professional growth through the mechanisms in our revised model in Figure 2. In the case, the self-assessment tool was bundled with other professional growth strategies: peer coaching, observation by external change agents, and focused input on teaching strategies.

Giving teachers access to a self-assessment tool without these supports will likely produce weaker outcomes. Without the consistency checks provided by feedback from peers or others, some teachers will generate inflated self-appraisals that reduce motivation to change. For example, a meta-analysis of 48 studies found that adult students rated their performance almost half a standard deviation higher than instructors (Falchikov & Boud, 1989).

Teachers who under-rate their performance or who accurately appraise themselves as low performers are also unlikely to change due to the depressing effect of negative self-assessment

on teacher efficacy. Teachers with low self-efficacy are less likely to implement new teaching ideas (evidence reviewed in Ross, 1998).

Finally, teachers who accurately self-appraise a need for change but have no support for implementing it are unlikely to be able do so, particularly in a domain such as mathematics education where change in practice has been found to be so difficult (Ross et al., 2002). For example, giving students regular feedback on performance without the means to improve has negative effects on their achievement (Fuchs, Fuchs, Karns, Hamlett, Katzaroff, & Dutka, 1997).

Earlier (Ross & Regan, 1993) we developed a theory of peer involvement in professional growth, which structured how curriculum consultants share professional experiences. We found evidence of professional growth when certain conditions were met: the consultant experienced dissonance between his/her current practice and a valued alternative; the teacher conceptualized how the alternative could be integrated into the core of his/her practice; the consultant experimented with the synthesis; the consultant consolidated the integration into routine operations. An in-service program developed for consultants based on this theory (Regan & Ross, 1991) had a significant impact on the quality of their strategies for bringing about change in schools (ES=.85) (Regan & Ross,1992). In retrospect we recognize similarities between the consultant study and our mathematics case. In the consultant in-service we provided a multi-dimensional rubric of consultant practice with levels of performance for each dimension, a structured self-assessment protocol, opportunities to share consultant strategies with peers, feedback from peers and researchers, and in-service on specific strategies for increasing school change. This retrospective confirmation of key elements of in-service design used in the mathematics teacher study strengthens the credibility of the findings from Barry's case and confirms the validity of our theory of how self-assessment contributes to professional growth.

The provision of a self-assessment tool based on well-defined teaching standards can strengthen in-service sessions that provide opportunities for teachers to observe one another and talk about classroom attempts to improve teaching.

References

Anderman, E. M., Anderman, L. H., & Griesinger, T. (1999). The relation of present and possible academic selves during early adolescence to grade point average and achievement goals. *Elementary School Journal, 100*(1), 3-18.

Arter, J., Spandel, V., Culham, R., & Pollard, J. (1994, April). *The impact of training students to be self-assessors of writing*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117-148.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education, 16*(239-253).

Cameron, J., & Pierce, W. (1994). Reinforcement, reward, and intrinsic motivation. *Review of Educational Research, 64*(3), 363-424.

Cancro, G. (1992). *The interrelationship of organizational climate, teacher self-efficacy, and perceived teacher autonomy*. Unpublished doctoral dissertation, Fordham University, New York, NY.

Chester, M., & Beaudin, B. (1996). Efficacy beliefs of newly hired teachers in urban schools. *American Educational Research Journal, 33*(1), 233-257.

Czerniak, C., & Schriver-Waldon, M. (1991, April). *A study of science teaching efficacy using qualitative and qauntitative research methods.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Lake Geneva, WI.

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59*(4), 395-430.

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C., Katzaroff, M., & Dutka, S. (1997). Effects of task-focused goals on low-achieving students with and without learning disabilities. *American Educational Research Journal, 34*(3), 513-543.

Goddard, R. D. (2001). Collective efficacy: A neglected construct in the study of schools and student achievement. *Journal of Educational Psychology, 93*(3), 467-476.

Goddard, R. D. (2002b). A theoretical and empirical analysis of the measurement of collective efficacy: The development of a short form. *Educational and Psychological Measurement, 62*(1), 97-110.

Goddard, R. D., Hoy, W. K., & Hoy, A. W. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Education Research Journal, 37*(2), 479-507.

Herman, P., Meece, J. L., & McCombs, B. (2000, April). *Teacher experience and teacher efficacy: Relations to student motivation and achievement.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Hickey, D. T., Kindfield, A. C. H., & Christie, M. A. T. (2003). Integrating curriculum, instruction, assessment, and evaluation in a technology-supported genetics learning environment. *American Educational Research Journal, 40*(2), 495-538.

Imants, J., & Van Zoelen, A. (1995). Teachers' sickness absence in primary schools, school climate and teachers' sense of efficacy. *School Organization, 15*(1), 77-86.

Kohler, F. W., Ezell, H. K., & Paluselli, M. (1999). Promoting changes in teachers' conduct of student pair activities: An examination of reciprocal peer coaching. *Journal of Special*

Education, 33(3), 154-165, 88.

Lepper, M. R., & Hodell, M. (1989). Intrinsic motivation in the classroom. In C. Ames, & R. Ames (Eds.), *Research on Motivation in Education (Volume 3)* (p. ??). San Diego: Academic Press.

Licklider, B. L. (1995). The effects of peer coaching cycles on teacher use of a complex teaching skill and teachers' sense of efficacy. *Journal of Personnel Evaluation in Education, 9*(1), 55-68.

Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experience in restructuring schools. *American Journal of Education, 106*(4), 532-575.

Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs.* San Francisco: Jossey-Bass.

Mascall , B. (2003). *Leaders helping teachers helping students: The role of transformational leaders in building teacher efficacy to improve student achievement.* Unpublished doctoral dissertation, University of Toronto, Toronto.

McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education, 10*(2), 209-220.

McDougall, D., Lawson, A., Ross, J., MacLellan, A., Kajander, A., & Scane, J. (2000). *A study on Impact Math implementation strategy for the Ontario Mathematics Curriculum Grades 7 & 8*. Toronto: OISE/UT.

McLymont, E. F., & da Costa, J. L. (1998, April). *Cognitive coaching the vehicle for*

professional development and teacher collaboration. Paper presented at the annual meeting of the American Educational Research Association, San Diego.

Miles, M., & Hubberman, A. (1994). *Qualitative data analysis: An expanded source book.* 2nd edition. Beverly Hills, CA: Sage.

Moore, W. P., & Esselman, M. E. (1994, April). *Exploring the context of teacher efficacy: The role of achievement and climate.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Moore, W. P., & Esselman, M. E. (1992, April 20-24). *Teacher efficacy, empowerment, and a focussed instructional climate: Does student achievement benefit?* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Morrison, G., Walker, D., Wakefield, P., & Solberg, S. (1994). Teacher preferences for collaborative relationships: Relationship to efficacy for teaching in prevention-related domains. *Psychology in the Schools, 31*, 221-231.

Muijs, D., & Reynolds, D. (2001). *Being or doing: The role of teacher behaviors and beliefs in school and teacher effectiveness in mathematics, a SEM analysis*. Paper presented at the the annual meeting of the American Educational Research Association, Seattle.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Raudenbush, S. W., Rowan, B., & Cheong, Y. (1992). Contextual effects on the self-perceived

efficacy of high school teachers. *Sociology of Education, 65*, 150-167.

Riordan, J. E., & Noyce, P. E. (2001). The impact of two Standards-Based curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education, 32*(4), 368-398.

Regan, E., & Ross, J. (1991). *Professional development activities for the reflective consultant*. Toronto: OISE Press.

Regan, E. M., & Ross, J. A. (1992-a). Professional development for school consultants: An evaluation. *Alberta Journal of Educational Research, 38*(3), 189-204.

Rosenholtz, S. J. (1989). *Teachers' workplace: The social organization of schools*. New York, NY: Longman.

Ross, J. A. (1998a). The antecedents and consequences of teacher efficacy. In J. Brophy (Ed.), *Research on Teaching. Vol. 7.* (pp. 49-74). Greenwich, CT: JAI Press.

Ross, J. A. (1992-b). Teacher efficacy and the effect of coaching on student achievement. *Canadian Journal of Education, 17*(1), 51-65.

Ross, J. A., & Cousins, J. B. (1993a). Enhancing secondary school students' acquisition of correlational reasoning skills. *Research in Science & Technological Education, 11*(3), 191-206.

Ross, J. A., Hogaboam-Gray, A., McDougall, D., & Bruce, C. (2001). The contribution of technology to the implementation of mathematics education reform: Case studies of grade 1-3 teaching . *Journal of Educational Computing Research, 26*(1), 123-140.

Ross, J. A., Hogaboam-Gray, A., McDougall, D., & Le Sage, A. (2003). A survey measuring implementation of mathematics education reform by elementary teachers. *Journal of Research in Mathematics Education, 34*(4), 344-363.

Ross, J. A., Hogaboam-Gray, A., & Rolheiser, C. (2002b). Student self-evaluation in grade 5-6 mathematics: Effects on problem solving achievement. *Educational Assessment, 8*(1), 43-58.

Ross, J. A., & McDougall, D. (2003). The development of education quality performance standards in grade 9-10 Mathematics teaching. Peterborough, ON: OISE/UT Trent Valley Centre.

Ross, J. A., McDougall, D., & Hogaboam-Gray, A. (2002). Research on Reform in Mathematics Education, 1993-2000. *Alberta Journal of Educational Research, 48*(2), 122-138.

Ross, J. A., & Regan, E. M. (1993a). Sharing professional experience: Its impact on professional development. *Teaching and Teacher Education, 9*(1), 91-106.

Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A. (1999). Effects of self-evaluation training on narrative writing. *Assessing Writing, 6*(1), 107-132.

Ross, J. A., & Starling, M. (in press). Effects of self-evaluation training on achievement and self-efficacy in a computer-supported learning environment. *Computers in Human Behavior*.

Schunk, D. H. (1981). Modelling and attributional effects on children's achievement: A self-efficacy analysis. *Journal of Educational Psychology, 73*(1), 93-105.

Senger, E. S. (1998). Examining teacher and classroom supports for student invention. *The Proceedings of the International Group for the Psychology of Mathematics Education, North America, 20*, 234-243.

Simon, M. A., & Tzur, R. (1999). Explicating the teacher's perspective from the researchers' perspectives: Generating accounts of mathematics teachers' practice. *Journal for Research in Mathematics Education, 30*(3), 252-264.

Soodak, L. C., & Podell, D. M. (1998). Teacher efficacy and the vulnerability of the difficult-to

teach student. In In J. Brophy (Ed.), *Advances in research on teaching. Vol. 7. Expectations in the classroom* (pp. 75-110). Greenwich, CT: JAI Press.

Stipek, D., Recchia, S., & McClintic, S. (1992). Self-evaluation in young children. *Monographs of the Society for Research in Child Development, 57*, 1-84.

Watson, S. (1991). *A study of the effects of teacher efficacy on the academic achievement of third-grade students in selected elementary schools in South Carolina*. Unpublished doctoral dissertation, South Carolina State College, Orangebury, SC. UMI 9230552.

Wineburg, M. S. (1995, April). *The process of peer coaching in the implementation of cooperative learning structures*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Wood, T., & Sellers, P. (1997). Deepening the analysis: A longitudinal assessment of a problem-based mathematics program. *Journal for Research in Mathematics Education, 28*, 163-186.

Yin, R. (2003). *Case study research: Design and methods. 3rd edition*. Newbury Park, CA: Sage.

Table 1 Excerpt from Standards-Based Mathematics Teaching Rubric

Dimension	**Level 1**	**Level 2**	**Level 3**	**Level 4**
9. Student Assessment: *Transparency*	The teacher's criteria for assessing student work are defined during marking.	The teacher defines the criteria for assessing student work before administering the assessment but does not disclose the criteria to the students beforehand.	The teacher defines the criteria for assessing student work and discloses the criteria and assessment procedures to students before administering the assessment.	The teacher negotiates with students the criteria for assessing student work and the assessment procedures before administering the assessment.

From Ross & McDougall (2003).

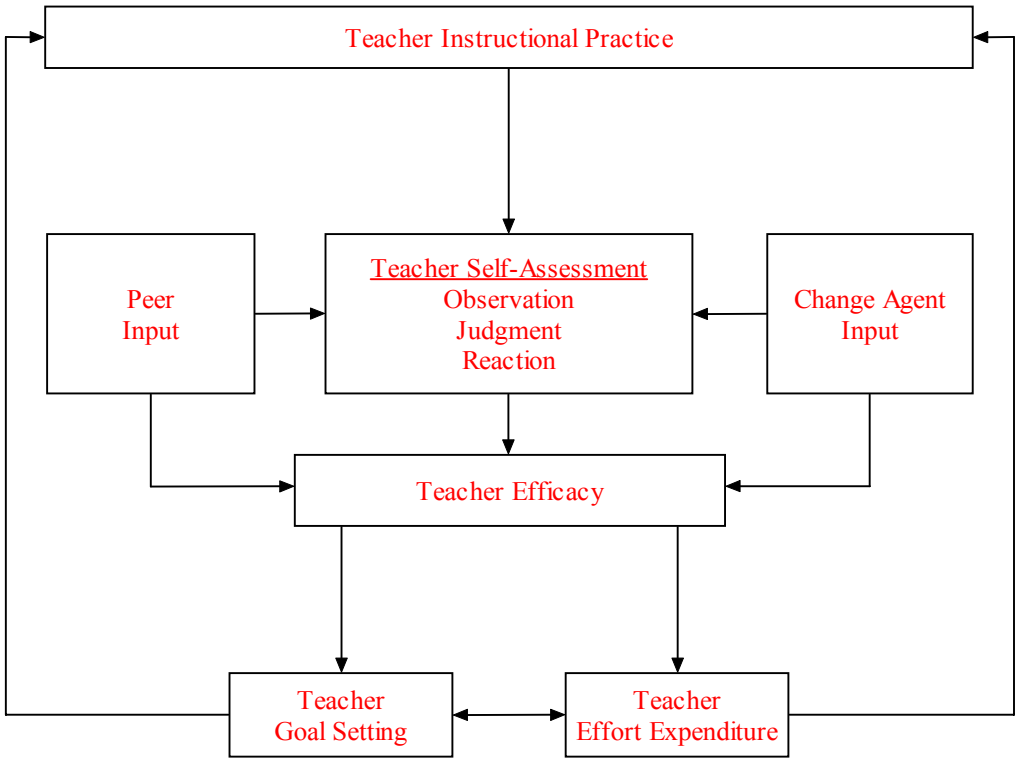

Figure 1: Initial Model of Teacher Self-assessment as a Mechanism for Teacher Change

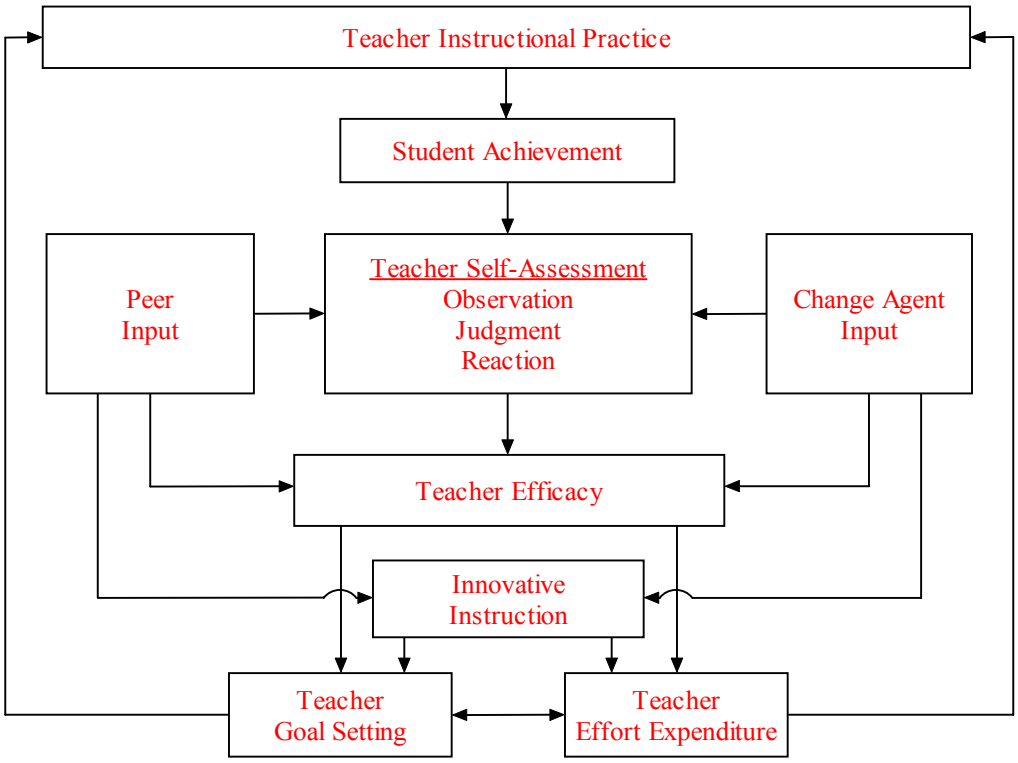

Figure 2: Refined Model of Teacher Self-assessment as a Mechanism for Teacher Change

Appendix: Observation Template and Coding Scheme

Example of Observation Template

Dimension 9: Student Assessment

Refers to how the teacher collects and interprets data on the quality of student performance. (may require discussion with teacher and / or students)

Guiding Questions

• What assessment strategies are observable in the classroom / in teacher-student interaction?
• Are students involved in the assessment design process?
• Are students involved in self and / or peer assessment?
• Are students aware of how they are being assessed?
• Are a variety of assessment strategies implemented?

Possible Evidence

- Evidence of rubrics in use (several levels of performance described)
- Teacher records observations during and / or after class
- Students use assessment tools such as portfolios, learning logs, journals
- Students provided with frameworks (verbal or written) to engage in self-assessment or peer assessment
- Students participate in negotiating evaluation schemes with the teacher

Observations

Coding Scheme

The coding scheme for the first two research questions (In what ways did Barry think he had changed? To what extent did our observations confirm or disconfirm Barry's claims about change?) was driven by our rubric; i.e., our theoretical framework for representing variation in mathematics teaching. Each chunk (e.g., interview utterance or observed vignette) was identified as one of the following 18 codes. Since the dimensions of our rubric overlap, we also assigned additional codes to the same chunk if the additional code was one of the target dimensions for that teacher. (D=rubric dimension)

- D1 Scope
- D2 Opportunity to Learn
- D3 Confidence
- D4a Tasks – Strategies
- D4b Tasks – Representations
- D5 Constructing Knowledge
- D6 Teacher's Role
- D7a Tools – Manipulative Use
- D7b Tools – Technology Use
- D8a Student Interaction – Instruction
- D8b Student Interaction – Tasks
- D8c Student Interaction – Communication
- D9a Assessment - Purpose
- D9b Assessment - Transparency
- D9c Assessment – Variety
- D10a Disciplinary Knowledge - Dynamic
- D10b Disciplinary Knowledge - Connections
- D11 – used for other important dimensions of mathematics programs that were not included in our rubric, e.g., classroom management

The coding scheme for the third research question (What factors contributed to and/or impeded teacher change?) was driven by the elements of our treatment design. The first seven factors identify elements of our intervention that might influence teacher practice (particularly a change). The last three factors identify variables other than the elements of our intervention that might influence teacher practice.

Intervention elements:

- TREAT the treatment as a whole, without differentiating elements of it
- PC peer conferencing
- RU rubric for implementing Standards-Based teaching
- SA self-assessment tools
- INF information about Standards-based teaching, i.e., workshops and interaction with individual researchers
- RES being observed or interviewed by the researcher

- OI other aspects of our intervention that were not part of our design, e.g., teachers might believe they were chosen for their excellence and now want to live up to it

Elements Unrelated to the Intervention:

- PERS personological factors such as the teacher's prior experience, family life
- SCHL school-related factors
- OTH other factors such as Ministry curriculum materials, district policy, etc.