2018

A Single Case Study on the Influence of Feedback on the Instructional Quality of a Preservice Elementary Teacher in Mathematics: The Instructional Quality Assessment Toolkit as a Resource

Makini Campbell Sutherland
University of Central Florida

A SINGLE CASE STUDY ON THE INFLUENCE OF FEEDBACK ON THE INSTRUCTIONAL QUALITY OF A PRESERVICE ELEMENTARY TEACHER IN MATHEMATICS: THE INSTRUCTIONAL QUALITY ASSESSMENT TOOL AS A RESOURCE

by

MAKINI CAMPBELL-SUTHERLAND

B.Ed University of the West Indies (Mona), 2008

MSc. Florida International University, 2010

A dissertation submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

in the College of Education and Human Performance

at the University of Central Florida

Orlando, Florida

Spring Term

2018

Major Professor: Dr. Juli K. Dixon

ABSTRACT

The purpose of this single case study was to determine if a preservice elementary teacher's instruction would be influenced by feedback on mathematics lessons. The focus of the research was on the use of the Instructional Quality Assessment (IQA) (Boston 2012) toolkit as an integral part in the feedback process. The IQA toolkit provides number ratings as well as qualitative descriptive ratings of various aspects of the mathematics lesson, defined under two constructs, or groups, labeled Academic Rigor and Accountable Talk.

The researcher evaluated a preservice teacher's instruction on three occasions and facilitated feedback sessions following those observation sessions, integrating the number ratings of the IQA toolkit with the qualitative descriptions of expectations. It was found that there was an increase in ratings for both constructs of Academic Rigor and Accountable Talk over the period of three feedback cycles. There was a difference in how the students were facilitated in instruction over the three observations. The teacher became more aware of some of her behaviors in the classroom that contributed to the type of instruction given to students.

In reviewing the literature, there was limited evidence of the use of the IQA toolkit for the iterative process of teaching, feedback, and teaching informed by feedback. This research is therefore useful in expanding the use of the IQA toolkit for feedback purposes. Preservice teachers as well as in-service teachers can benefit from feedback focused on mathematics teaching that makes them more aware of their strengths and weaknesses so they are able to adjust instruction based on the feedback received.

ACKNOWLEDGMENTS

The process to complete my dissertation was interesting and met with some eventful situations. To complete, I could not have done it without the support of some individuals. First to the Almighty God who gave me the strength during those difficult moments. To my committee chair Dr. Juli Dixon for her guidance, encouragement, and unwavering support through my challenges. To the rest of my committee Dr. Janet Andreasen, Dr. Kristopher Childs, Dr. Elizabeth Hoffman, and Dr. Farshid Safi, thank you for your feedback and support. To my fellow doctoral students Aline and Brianna, I appreciate your support throughout the process. Finally I thank my amazing husband Andre, our son Timothy-Cole, my mother Judith and dad Selvin. You are my inspiration.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

In an era where the performance of students is of utmost importance and U.S. students continue to lag behind their peers internationally (Ball, 2003; Carnoy & Rothstein, 2013), the focus on instruction and curricular reform is at the forefront of the minds of stakeholders and other interests. Teacher preparation, among many other aspects, has gained the attention of decision makers and no doubt contributes to current reform efforts for teacher education (Darling-Hammond, 2014). While there are numerous teacher preparation programs, executed through various means and modalities, the performance of teachers in mathematics at the elementary level continues to garner attention and researchers of mathematics have stated that teachers are not prepared for the mathematics they are required to teach (Ball, 1990; Ryan & Williams, 2007; Isiksal & Cakiroglu, 2011; Maher & Muir, 2013).

Strengthening teacher education, then, is increasingly a concern for policymakers in the United States (Darling-Hammond, 2014) and therefore the experiences of the preservice teacher ought to be considered as an important part to achieving this goal. The teaching internship experience can have a significant impact on the types of teachers interns will become (Rhoads, Radu, & Weber, 2010) and the supervisors' success in preparing the interns can be measured by how well students are able to self-evaluate to result in improved classroom performance (O'Shea, Hoover, & Carroll, 1988). Success for interns involves conferences with the supervisor that demonstrate effective conferencing techniques in an effort that the intern develops trust and cooperates well with the supervisor (O'Shea, Hoover, & Carroll, 1988).

Evaluation of instruction often takes place in education. The No Child Left Behind Act (No Child Left Behind [NCLB], 2001) focused heavily on the evaluation of teachers through the

use of student achievement, often evidenced by standardized testing results. However, the Every Student Succeeds Act (2015) provides stakeholders in the United States the opportunity to lessen their focus on teacher evaluation tied to student outcomes (Klein, 2016). Practicing teachers are evaluated in their classrooms and so it makes sense that preservice teachers (PSTs) are evaluated on their teaching experiences in the classroom as well. Strengthening the clinical practice is one of the more important aspects for improving the competence of new teachers and recent efforts have been focused on the importance of well-supervised clinical practice (Darling-Hammond, 2014). If teacher education programs are to be successful in preparing high quality teachers who are able to adequately deal with the complexities of classrooms, the evaluation of comprehensive field preparation is a critical component (Wilson, 1996). However, the nature of clinical supervision does not necessarily demonstrate this importance since the task is usually left to people who are available and not necessarily the ones who possess specialized knowledge (Meade, 1991). In mathematics, this specialized knowledge should consist of both Specialized Content Knowledge (SCK) and Knowledge of Content and Teaching (KCT). SCK is mathematical knowledge beyond what is expected of someone who is educated in general. KCT includes knowing mathematics as well as knowing how to teach it (Ball, Thames, & Phelps, 2008). Regardless of who is conducting the supervision, observation of instruction is a main aspect of the process and evaluation tools are often used to assist in accomplishing the task.

The tools that are used in evaluation vary depending on factors such as state, district, and institutional requirements. While districts in different states use their choice of evaluation tools, current evaluation tools such as those by Danielson (2007) and Marzano (2011) are commonly used to evaluate teachers. According to the 2015 report from the Florida Department of Education, during the 2013-2014 school year, more than 50% of schools in Florida reported to

have used Marzano or Danielson evaluation tools for teacher evaluation (Florida Department of Education, 2015). While these tools are commonly used, they are not specifically geared towards assessing mathematics lessons with the exception of a new framework being developed by the Danielson Group to include mathematics evaluation specifically, now in the initial draft phase (www.danielsongroup.org). Boston, Bostic, Leeseig, and Sherman (2015) stated that evaluation tools specific to mathematics are able to provide information on aspects of practice that are important in students' learning of mathematics, provide a concrete structure for new practices to be developed, and identify standards that are important for implementation that have been supported by research. Although Marzano's and Danielson's tools are used widely, tools specific to mathematics are also used in the evaluation of mathematics instruction. Teacher evaluation tools specific to mathematics such as the Instructional Quality Assessment (IAQ) toolkit have been used by researchers to evaluate mathematics lessons specifically. The IQA toolkit (Appendix A) is a set of rubrics designed to measure the quality of mathematics instruction using both statistical and descriptive components (Boston et al., 2015).

Within the IQA toolkit, observation of instruction is categorized into two main groups also called constructs, of Academic Rigor and Accountable Talk, and include sub-constructs of specific areas to be measured during assessment of instruction from the task and whole class discussions capturing the Potential of the Task, the quality of students' discussion, mathematics connections made by the teacher, and student participation. The numerical rating determined by applying the constructs of the IQA toolkit is used to identify the quality of aspects of instruction. Therefore specific areas of weakness can be identified and be addressed (Boston, 2012; Matsumura et al, 2006). While there is discussion surrounding use of the IQA toolkit in

formative assessment that would involve feedback, (Junker et al., 2006; Boston et al., 2015), most of the research related to the use of the IQA toolkit does not address this facet of its use. Feedback using the IQA toolkit could enable PSTs to plan and make adjustments based on identified areas of weakness prior to being reassessed. The teacher education process should allow for PSTs to become experts in making sense of the theory they learn and transferring that into practice to meet the needs of students (Darling-Hammond, 2014). If mathematics-specific evaluation tools are used with feedback that focuses on areas specific to mathematics development in the instructional process, it is possible for PSTs to transfer the theory they learned through coursework to practice using an iterative process of teaching – feedback – teaching informed by feedback, so that the needs of students can be met.

What follows is a description of a research study that examined the influence of feedback on the instructional practices of a final year preservice elementary teacher in mathematics. The constructs of the IQA toolkit were used as a guide for focused feedback on specific descriptive criteria associated with the tool. This study discusses ways in which a PST was influenced by the feedback process and how this influence was reflected in her instructional experience from planning to reflection.

Statement of the Problem

There are many teachers who graduate from teacher training programs each year poised to become a part of the core of teachers needed in the classroom. However, researchers continue to report how underprepared teachers are to teach mathematics at the elementary levels, not possessing the kinds of knowledge needed to teach students conceptually (Isiksal & Cakiroglu, 2011; Maher & Muir, 2013). The pedagogical content knowledge of the PST as described by

Shulman (1986) should go beyond knowledge of the subject itself and tap into subject matter knowledge for teaching. The nature of this research was to assist a PST in developing this knowledge through the use of an iterative process of observing teaching and providing feedback regarding her teaching practice so that the PST is able to meaningfully engage her students in mathematics learning.

Rationale

Student achievement that falls below satisfactory levels raises the concern of the quality of mathematics instruction that students experience (Ball, 2003). The underlying assumption is that the discrepancy, or differences, in the level of student achievement is caused by differences in teacher effectiveness (Phillips, 2010). If the quality of instruction, tied to the teacher, is low, student achievement may also be affected, and could partially explain the achievement scores. The improvement of mathematics education is considered through the lens of a preservice teacher whose role, among other responsibilities, was to implement tasks that supported student engagement in reasoning and to facilitate problem solving and meaningful discourse. High quality instruction is process focused, considering the concepts involved, where learning is based on previous knowledge used to build new knowledge rather than answer focused (Kilpatrick, 2001). Discourse is a main element in the lesson implementation process, and is likely to foster the development of students through the eight mathematical practices (National Council of Teachers of Mathematics (NCTM), 2014). Teachers entering the profession are expected to possess the ability to engage students in these eight Standards for Mathematical Practice:

- make sense of problems and persevere in solving them
- reason abstractly and quantitatively

- construct viable arguments and critique the reasoning of others
- model with mathematics
- use appropriate tools strategically
- attend to precision
- look for and make use of structure
- look for and express regularity in repeated reasoning (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010, p. 6).

This study examined the implementation of feedback that is directed and purposeful using the IQA toolkit as an aid. A number of research projects using the IQA toolkit have been conducted, however, the element of feedback has not been extensively explored. Therefore, while there have been favorable results with the IQA toolkit as it relates to measuring instructional quality, research has not focused on using the IQA toolkit to determine its possible influence on teachers' instructional quality, through the process of feedback. Improving the instructional quality and experiences of PSTs allows them to be better prepared for their potential future roles as teachers in the classroom and that student performance to which their teaching is linked is likely to be improved. The ability to effectively engage students in the mathematical practices is not innately possessed but should be taught to PSTs as suggested in the study reported by Graybeal (2013), where pre-service teachers needed assistance to understand, identify, and collect evidence of the standards being enacted in classroom settings. Pedagogical content knowledge (PCK) through the work of Shulman (1986) identifies the intersection between content and teaching. PCK includes understanding what it takes for students to learn

mathematics, students' preconceptions, misconceptions, and knowledge of strategies that are likely to help students develop an understanding of concepts. PSTs then should be engaged in understanding the content as well as how to transfer that knowledge to support learning (McDonnough & Matkins, 2010).

The research questions that were analyzed in this single case study sought to assess ways in which feedback using the two main constructs of the IQA toolkit influenced the instructional quality of a PST.

Main Question: How does feedback using the Instructional Quality Assessment (IQA) toolkit influence the mathematics instructional practices of a final year preservice elementary teacher?

Sub-questions

1. In what ways does feedback using the construct of Academic Rigor as measured by the Instructional Quality Assessment toolkit influence the mathematics instructional quality of a final year preservice elementary teacher?
2. In what ways does feedback using the construct of Accountable Talk as measured by the Instructional Quality Assessment toolkit influence the mathematics instructional quality of a final year preservice elementary teacher?

Significance of the Study

It is hoped that this single case study will contribute to the literature that already exists on feedback in instruction. Specifically, it is hoped that the way feedback is provided to preservice teachers regarding their instructional quality in mathematics with the aim of making the instructional experiences for both students and teachers better, will be improved. This study explored a means to be able to provide focused feedback through the use of the IQA toolkit. The IQA toolkit was used as a guide to what is expected of PSTs. Although there has been

discussion surrounding the use of the IQA toolkit in professional development (Junker & Matsumura, 2006) research on feedback using the IQA toolkit, especially for PSTs of elementary mathematics has not been published. Additional studies that focus on the nature of the feedback that may influence instructional quality in a positive way are needed. Studies of this nature may inform teacher preparation, therefore equipping new teachers with the necessary skills in mathematics teaching that foster students to be critical thinkers and where the PST is able to engage students in the Standards for Mathematical Practice as outlined in the Common Core State Standards for Mathematics (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010)

Organization of the Dissertation

This dissertation is organized into five chapters. Chapter one includes an introduction, which reviews the problem, background of the study, the statement of the problem, the rationale for the study, and the significance of the study. Chapter two contains a review of relevant literature and the theoretical framework for the study. Chapter three details the research questions, methodology, and procedures. Chapter four presents the data and subsequent analyses and chapter five concludes with the discussion of the dissertation presenting the meaning of the findings, implications for future research, and the limitations of the study.

CHAPTER 2: LITERATURE REVIEW

The purpose of this study was to determine if feedback on instruction with the use of the Instructional Quality Assessment toolkit influenced the instructional quality of a preservice elementary teacher. This literature review describes studies on measuring instructional quality among teachers. Although these studies examined instructional quality through the Instructional Quality Assessment (IQA) toolkit, most did not delve into incorporating feedback and follow-up as an element in the research. Research on using the IQA toolkit in elementary mathematics classroom settings is limited with few studies having reported its use (Boston & Wolf, 2006; Junker et al., 2006; Quint et al., 2007) and even fewer focusing on PSTs specifically (Junker et al., 2006). As such this study provides further insight into an elementary mathematics PSTs final-year internship experience and her demonstrated instructional quality in facilitating mathematics lessons.

Additional literature on classroom evaluation tools commonly used to measure instructional quality will be discussed along with research outcomes associated with them. Earlier studies in measuring instructional quality in general focused on in-service teachers and on explaining statistical data. Since the literature on PSTs' instructional quality using the IQA toolkit is limited, it follows that the literature on feedback using the IQA toolkit is too, and therefore research on providing feedback and reassessing PSTs and analyzing the influence that such feedback would have on PSTs in a non-statistical manner is indeed useful. Analyzing problems in this way facilitates creating meaning through a theoretical lens, facilitating understanding of issues evident in their natural settings (Creswell, 2013).

Instrumental to this review of literature is the theoretical framework of the Feedback Intervention Theory used in the analysis of the data for this study. The description of the theory provides an understanding of its relevance to this research and how the processes described informed the findings that resulted from this study.

Building on the analysis of prior research and scholarship, in this chapter I argue that the use of the IQA toolkit as an element in the feedback process is a possible consideration for creating a positive influence on the feedback process of final year PSTs during their internship experience. Additionally, a critique of empirical investigations available will be provided highlighting the need for the study.

Constructs of Interest

In this research there are ideas that are presented as important themes from the IQA toolkit referred to as constructs. The constructs from the IQA toolkit important to this research are Academic Rigor and Accountable Talk. An additional important theme throughout this research is feedback. In light of the information that will be emphasized in the review of literature, it is important that the constructs of interest be defined and explained so there is a better understanding of the information that follows.

Academic Rigor. Mathematics teachers are expected to provide students opportunities to engage in rigor while learning the content. Rigor facilitates students to make connections, solve challenging problems, and explain and justify their solutions (Hull, Miles, & Balka, 2014). Academic Rigor is based on the idea that students should be exposed to and engaged in rich knowledge that is developed around mastery of content (Boston & Wolf, 2005).

The theoretical connection to Academic Rigor is provided by Stein and colleagues in the Mathematical Tasks Framework. Stein and Smith (1998) posit four levels of mathematical tasks with levels one and two that include memorization and procedures without connections described as low-level tasks. High-level tasks from levels three and four include procedures with connections and doing mathematics. The work of Stein and Smith (1998) is important with respect to the IQA toolkit because the extent to which students are provided with opportunities to learn mathematics with understanding is essential to measure the Academic Rigor of mathematics instruction (Boston et al., 2015). The cognitive demand of the task students are given can be a fundamental indicator of the rigor of the instruction (Boston & Wolf, 2005). Students engaged in Academic Rigor should be provided with opportunities to learn mathematics conceptually by being supported to solve problems; develop, explain, and justify solutions; validate their own solutions, learning from previous efforts and remaining focused; and communicate ideas. The concept of rigor in mathematics is concerned with engaging students in learning mathematics conceptually (Boston & Wolf, 2006; Hull et al., 2014). In mathematics it means simply that students are given the opportunity to learn mathematics with understanding (Boston & Wolf, 2005).

Rigor in mathematics as explained by Hull et al. (2014) sees shifts from what was the norm in mathematics classes with rapidly paced classes, where many problems were completed, more homework assigned, and students worked independently and under pressure. The shift is geared towards an atmosphere that encourages students to explore collaboratively, in a supportive and encouraging environment, where homework is a natural and interesting extension and the pace is slower, yet students are scaffolded to push deeper. Essential to Academic Rigor is the facilitation of discourse within the mathematics classroom.

Accountable Talk. Discourse in mathematics relates to mathematical talk (Imm & Stylianou, 2012). Accountable Talk focuses on the quality of classroom discussions or classroom discourse with respect to the learning community and mathematics (Boston et al., 2015). Accountable Talk also refers to the talk that occurs in the mathematics classroom that contributes to learning in the classroom where students' views, rather that the views of the teacher, become the focus of discussions (Ittigson, 2002). In classrooms where Accountable Talk is encouraged, students are expected to listen to their classmates, ask questions of them to better understand their thinking, and justify their thinking using evidence from their work (Ittigson, 2002)

Facilitating Accountable Talk. Some observations may be made regarding facilitating Accountable Talk in mathematics classrooms. Student participation is central as students are called on to present ideas that initiate inquiry and challenge the ideas presented by their peers (Nathan & Knuth, 2010), similar to the third Standard for Mathematical Practice where students are expected to construct viable arguments and critique the reasoning of others (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). Mathematical practice three embodies the idea that students are engaged to build mathematically sound arguments by logically analyzing mathematical statements. Students are also encouraged to listen to their peers and make decisions about the explanations and justifications of their arguments and use questioning to clarify the explanations of their peers. The teacher's role is concerned with eliciting and engaging students' thinking and knowing when to step in or step aside, giving students the space they need to explore, and allowing students to create their own ideas (Nathan & Knuth, 2010).

A critical component to facilitating Accountable Talk is the inclusion of problems or tasks that engage and challenge students, rather than focusing on problems used exclusively to help students select operations (Ittigson, 2002). The task of posing problems that encourage thinking and discussion and the use of multiple solution strategies is of importance (Ittigson, 2002) and suggests that the role of the teacher supersedes acknowledging students' approaches to solving predictable tasks as being correct or not (Stein et al, 2008). Tasks that ask students to perform a routine that has been memorized limits the opportunity for students to think conceptually and make the appropriate mathematical connections, compared to tasks that create opportunities for students to think (Stein & Smith, 1998).

Central to rich discourse is the type of questions that teachers ask, but teachers often face the challenge of orchestrating discussions that use students' responses to advance mathematical learning (Stein, Engle, Smith, & Hughes, 2008). Teacher questioning is a critical aspect of teachers' work (Boaler & Broadie, 2004) but also requires pedagogical content knowledge of the teacher to be successfully included in lessons (Shulman, 1986). The questions asked by the teacher must take into consideration the context of the instruction in relation to the tasks students are given (Herbert & Wearne, 1993) and when faced with a myriad of responses from students, the teacher must find ways to help them gain a deeper understanding of significant aspects of the lesson (Stein et al., 2008).

Stein and colleagues (2008) posit five practices that may be adopted in facilitating mathematical discussions around tasks that are challenging. These include anticipating students' responses to these tasks, monitoring students' responses during the individual work period where they are allowed to explore, selecting particular students to present their mathematical responses during the discuss and summarize phase, purposefully sequencing students' responses to be

displayed for the whole group, and helping the class to make the mathematical connections between and among the responses of their peers. The process suggested by Stein and colleagues is aligned with their suggestion that teachers should carefully plan, thereby anticipating students' possible responses and making informed instructional decisions, using knowledge of students' current mathematical thinking. For PSTs to be able to do this, the feedback they receive on their instruction should be purposeful.

Feedback. To improve the performance of PSTs in the internship process, feedback is considered an important aspect in the development of intern experiences, providing opportunities for them to retry and continue to improve (Darling-Hammond, 2014). Khachatryan (2015) discusses the need for qualitative feedback in that teachers often do not understand the evaluation process because feedback summed up in a numerical value does not provide details of performance. She goes on further to explain that teachers who are evaluated by individuals who are not knowledgeable of specific content areas are placed at a disadvantage. Improving teaching involves providing timely feedback after assessment and engaging teachers to reflect on the feedback data (Darling- Hammond, 2014). The assessor who is not knowledgeable in the critical areas would likely have difficulty in providing the type of feedback most likely to benefit the individual being assessed**.**

Types of Feedback. Khachatryan (2015) cited three types of feedback, self-feedback, product feedback, and process feedback. Self-feedback directs the recipient's attention to him or herself as the teacher, with comments focused on the teacher's skills and dispositions and will not necessarily improve the performance of the teacher because of the personal nature of the feedback. Feedback is better directed to the learning processes that take place during instruction. Increased motivation in the teacher must be achieved for it to be considered useful to impact

future performance. Feedback should focus on the product or the end result of performance; this is referred to as product feedback. Product feedback takes an evaluative form that focuses on how well the recipient accomplished the task, whether positive or negative. This type of feedback has the potential to increase motivation and its effect on the teacher's performance. Khachatryan's research focused on the type of feedback that results in improved practice among teachers and stated that if teachers receive self-feedback their teaching is unlikely to change. The author also states that feedback focused on product or outcome will increase motivation. However, if teachers receive process feedback, then they are able to learn about specific practices of themselves during the lesson and the teacher is therefore better able to modify his or her practice as needed.

Khachatryan's research involved an administrator and four high school teachers as participants in four subject areas of mathematics, history, social science, and science. In total the four teachers received 259 pieces of feedback. Khachatryan examined this feedback and found that 45% (N=117) of the feedback outlined the processes or steps taken by the teacher during the instruction phase (process) and 39% (N=68) of the feedback was more evaluative and focused on the product giving attention to motivational processes. Responses from the teacher participants were varied, with three main types of responses, (1) those who were positive regarding the feedback they received and felt that it validated their teaching practices, (2) those who thought that improvements could be made, and (3) others who questioned the meaning and accuracy of the feedback comments. While this study stated that rubrics were used, the nature of the rubrics were not discussed beyond that they aligned to California standards for the four subject areas assessed.

The results showed that close to half of the comments of the assessor were detailed and provided relevant information in the feedback. The major area of focus for the feedback was on the process and product and less about self. The feedback was deemed effective since the focus was on the type of feedback that produced increased effects on future performance (Kluger & DeNisi, 1996). Kluger and DeNisi in the explanation of their theory state that effective feedback intervention is followed by an action that seeks to decrease the discrepancy that is identified.

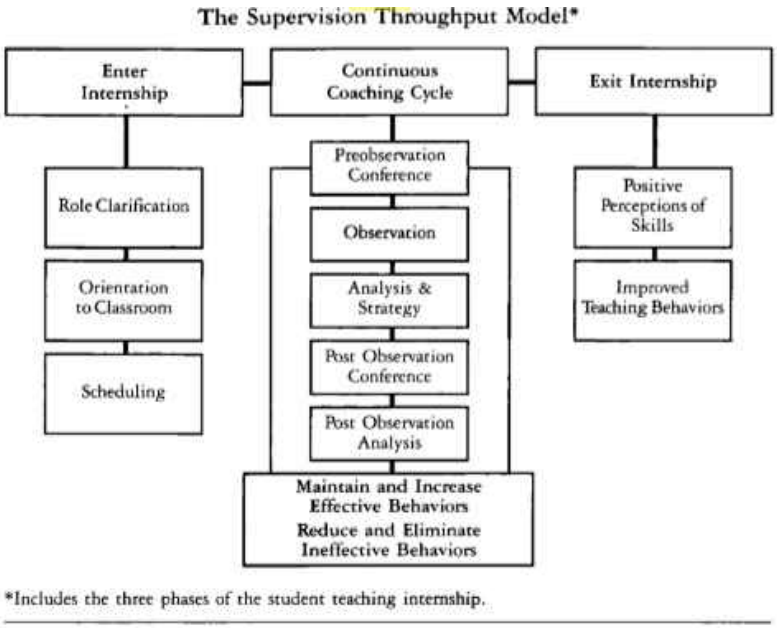
Anderson and Radencich (2001) conducted a study on feedback with preservice teachers. The sources of feedback for the study were peers, teachers, and the university supervisors. The study consisted of 34 elementary education preservice teachers in their final year of training. Qualitative data sources included data forms, dialogue journals, and course evaluations. The preservice teachers indicated they valued feedback from all the sources but especially from the classroom teacher and the university supervisor. Important to note is that the university supervisor assumed the role of a coach rather than an evaluator. The researchers stated they directed the focus of the university professors on providing the environment where preservice teachers learn how to teach, rather on a focus on what to teach.

The role of the consultant in the study by Capizzi, Wehby, and Sandmel (2010) was similar to that of Anderson and Redencich. Capizzi et al. in their study explored consultation and self-evaluation to improve the instructional delivery of three preservice special education teachers. In this study, the lessons of the participants were videotaped. The lesson took the traditional form and was assessed on components such as the introduction to the lesson, guided practice, modeling of instruction, independent practice, review, and closure. The participants met with a consultant who evaluated the components of the instruction and provided feedback. It

was found from the research that the consultation improved the number of lesson components considered for all participants.

O'Shea, Hoover, and Carroll (1988) discussed the Supervision Throughput Model (STM) as illustrated in figure 1. The STM contains three phases that they posit PSTs will go through during the internship process as they begin to develop practical skills related to teaching. The first phase as they enter internship focuses on the development of a relationship between the intern and the supervisor, where roles are clarified, PSTs become familiar with the classroom setting and sort out issues related to scheduling.

The next step, the Continuous Coaching Cycle (CCC), includes pre-observation conference, observation, analysis and strategy, post observation conference, and post observation analysis. Post observation conference is considered the most important aspect of the assessment process because of its reoccurrence throughout internship observations and its likely influence on behavior adjustments related to further observations. The post observation conference should focus on the behavior of the intern, patterns that have been observed, and feedback regarding teaching performance. The researchers caution that the supervisor should try to withhold value judgment so that an atmosphere is created where the intern is able to be less defensive and more productive in identifying problems and devising solutions. Interns will be able to evaluate themselves more effectively as they begin to identify their own problems and devise solution pathways to improve their craft.



Reprinted from "Effective Intern Conferencing," by L.J. O'Shea, N.L. Hoover, and R.G. Carroll, 1988, Journal of Teacher Education, p.18. Copyright 1988 by SAGE publications

Figure 1: The Supervision Throughput Model (STM).

The STM demonstrates possible processes the PST will go through in their internship experience. As PSTs exit internship the STM suggests that PSTs will demonstrate positive perception of skills as well as improved teaching behaviors and demonstrate the knowledge needed to have a successful internship. The skills interns need are varied but central to their success is the knowledge they need as teachers of mathematics.

Knowledge mathematics teachers need

The Association for Childhood Education International (ACEI) *Elementary Education Standards* (2007) state that teachers need to know, understand, and use the major concepts and procedures that define the various mathematics content areas and that students should be engaged continuously in skills such as reasoning, problem solving, proof, and making

connections. These standards conceptualize what teachers who teach mathematics should know and be able to do. Teachers need to have a thorough understanding of the mathematics they are required to teach as well as explanations of mathematical ideas that are accurate (Ball, 1990; Ma, 1999). ACEI provides standards similar to that of the National Council for Teachers of Mathematics (NCTM, 2014). These standards state that quality teaching includes engaging students in meaningful discourse, allowing students to explain and reflect on their thinking and that of their classmates, problem solve, ask questions, and be engaged in understanding the distinctions of each topic while connecting to other areas in mathematics.

The previous statements explain clearly what Shulman (1986) embodied in his seminal piece on teacher content knowledge. Shulman identified subject matter knowledge and pedagogical content knowledge (PCK) as aspects that are important to mathematics teaching. He discussed that the mathematics teacher should possess content knowledge specific to mathematics and be skilled in the teaching of mathematics to identify students' errors and misconceptions, understand their difficulty levels, and implement instruction that is developmentally appropriate for students (Boston et al., 2015; Ball, Thames, & Phelps, 2008; Cochran, King, & DeRuiter, 1991).

Ideally PSTs should possess the knowledge that is required. However, many graduate from teacher education programs without the level of conceptual understanding in mathematics needed to teach at the elementary level (Ball, Lubienski, & Mewborn, 2001). Research on PSTs identify that they may not be fully prepared to teach the mathematics concepts required (Ball, 1990; Ryan & Williams, 2007; Isiksal & Cakiroglu, 2011; Maher & Muir, 2013). PSTs have difficulty with mathematics content and pedagogy since the skill of teaching is developed over time as topics are taught repeatedly (Ball, Lubienski, & Mewborn, 2001).

Maher and Muir (2013) found that incorrect solutions to tasks of elementary PSTs (N=20) included procedural as well as computational errors when attempting to solve problems using standard algorithms. Similarly, Isiksal and Cakiroglu (2010) concluded that preservice elementary teachers (N=17) had computational and procedural errors solving fraction problems contributing to the idea from other researchers that preservice teachers as well as beginning teachers are not equipped to fully teach mathematics at levels that are required.

The knowledge that PSTs should possess will be called upon for use in their field experiences to plan and execute lessons. In the initial field experience of PSTs, interaction with students and teachers are at a minimum since the internship consists mainly of teacher observation. PSTs are rarely asked to teach lessons extensively during their initial internship or field experience (Hoyt & Terantino, 2015). However, the final internship field experience is varied but takes the form of PSTs teaching and conducting lessons. For some teacher preparation programs, students participate in a single internship.

There is a need to ensure that PSTs have authentic experiences to reflect the classrooms of today (Council for the Accreditation of Educator Preparation [CAEP], 2013). However, many programs do not possess the aspects of clinical and didactic curriculum needed in effective preparation programs (Darling-Hammond, 2010). Additionally, cooperating teachers with whom students work during field observations often know very little about the courses that students take and the people who teach in their teacher preparation programs are not always aware of the daily challenges that take place in the classroom (Zeichner, 2010). This disconnect could possibly lead to other problems related to the transition the PST has to make upon entering the classroom. As such research to minimize the potential difficulties PSTs may face is useful.

While the internship or practicum experience can be varied for each PST because of specific program differences or university requirements, PSTs are expected to use what they learn from their teacher education coursework in the classroom. The challenge occurs when they enter the field to find expectations different from what they learned. For coursework to have meaning, the in-the-moment demands and responses required of and by PSTs must be considered. PSTs will need to have the experience to redesign lessons because pre-packaged lessons are not always successful with the students that need to be engaged or are not focused on developing conceptual understanding (Darling-Hammond, 2010).

The internship process, then, should be carefully planned and directed especially for mathematics, considering the research that states that PSTs are not adequately prepared for the classroom. The research asserts that teachers have difficulties in mathematics and are not necessarily prepared by the time they go into the classrooms (Ball, 1990; Ryan & Williams, 2007; Isiksal & Cakiroglu, 2011; Maher & Muir, 2013). There is a need to support PSTs to deliver lessons considering the important aspects of instruction to engage students meaningfully. What follows is a review of classroom evaluation tools, with special attention given to the Instructional Quality Assessment (IQA) toolkit designed to measure the quality of mathematics instruction and can therefore be used as a tool for feedback during the internship process.

Classroom evaluation

Improving methods to evaluate teacher performance has been a focus of recent reform efforts. While there is much talk about standardized testing which reflects the effects of classroom instruction, classroom observations may be able to capture aspects of teacher performance that are important but may not be reflected in the standardized scores of students

(Steinberg & Sartain, 2015). Various frameworks are used throughout the school system for teacher evaluation. Two of the frameworks that are commonly used are by Danielson (2007) and Marzano (2011). Danielson (2007) sought to develop a framework that considered the complex nature of teaching encompassing planning, preparation, reflection, and interaction with colleagues. Student teachers are able to use frameworks such as Danielson's to reflect on their performance with more directed focus on their strengths and weaknesses. This facilitates them to be less focused on asking if a skill was observed but rather to consider how they have changed or developed in their craft over time (Roegman, Goodwin, Reed, & Scott-McLaughlin, 2015). Danielson (2007) provides details of the domains of her framework. The domains and accompanying competencies as shown in table 1 include planning and preparation, the classroom environment, instruction, and professional responsibilities. Each competency level is either unsatisfactory for level one, basic for level two, proficient for level three, or distinguished for level four.

Table 1: Domains and competencies of Danielson's (2007) framework

Domain	Competencies of the domain
Domain 1: Planning and preparation	• Demonstrating knowledge of content and pedagogy • Demonstrating knowledge of students • Setting instructional outcomes • Demonstrating knowledge of resources • Designing coherent instruction • Designing student assessments
Domain 2: The classroom environment	• Creating an environment of respect and rapport • Establishing a culture for learning • Managing classroom procedures • Managing student behavior • Organizing physical space
Domain 3: Instruction	• Communicating with students • Using questioning and discussion techniques • Engaging students in learning • Using assessment in instruction • Demonstrating flexibility and responsiveness
Domain 4: Professional responsibilities	• Reflecting on teaching • Maintaining accurate records • Communicating with families • Participating in the professional community • Growing and developing professionally • Showing professionalism

Research including the use of Danielson's framework is varied. Martin and Mertl (2014) report on research which sought to answer questions including one of determining if the

instrument was able to measure areas effectively based on the expectations of the Common Core State Standards for English Language Arts, including social studies, science, and mathematics literacy standards. Four hundred thirteen respondents among which were school administrators, teachers, and district administrators participated in the research. The findings were divided into 10 sections. Eighty-one percent of the respondents stated that they wanted an evaluation system that was aligned with the Common Core State Standards and that alignment would benefit both teachers (86% agreed) and administrators (88% agreed). Additionally 91% of the respondents felt that the framework was effective in evaluating overall teaching practice. Interestingly, 87% of the respondents felt that the framework still needed to change to better reflect the new standards. Practitioners were also reported to prefer one tool that would be used for all subject areas because it would be less cumbersome. However when participants were asked if they would prefer the framework to be customizable based on subject area, a majority of 70% agreed. The results from Martin and Mertl imply that although evaluation tools are useful and would be easier for the evaluator if only one tool was used, results would be more useful to teachers if they were based on a more subject-specific evaluation. It seems that subject specific evaluation tools are more valuable to those who are evaluated since those tools can help to focus on key features of intervention and treatment (Boston et al., 2015)

Steinberg and Sartain (2015) conducted a study in Chicago Public Schools focused on the causal impact on school performance of an evaluation system that was based on highly structured observations. Chicago teachers were previously evaluated based on a checklist of 19 classroom practices with the options of “strength”, “weakness,” and “does not apply”. The checklist approach was not well received by teachers and administrators alike and teachers who

were considered high performing criticized the ability of that system to provide meaningful feedback on their instruction. In fact Steinberg and Sartain reported less than 40% of principals felt that the checklist was adequate to address teacher underperformance. Steinberg and Sartain explained that the process moved from simple observation to a process that involved pre-observation where teachers were able to have discussions with the administrator, in which the rubric was reviewed and important information about the classroom discussed. The principal would then take detailed notes during the lesson about the actions of both students and teachers and the notes were matched with Danielson's framework rubric to rate teachers' performance on the 10 areas of instructional practice.

The intervention took place over a two-year period with two cohorts. Cohort one started in the first year and cohort two joined the second year. Two areas of the Danielson framework, classroom environment and instruction were the focus, with ratings of teachers as unsatisfactory, basic, proficient, and distinguished. Principals would meet teachers within a week of their observation for the post-observation conference. Evidence from the observation as well as the connection to the Danielson Framework would be discussed. While the way in which teachers were evaluated had shifted, the results from the intervention in the first year showed statistically significant gains in student achievement in reading but not mathematics. In the second year, when cohort two was added to the intervention, gains again were determined for reading but with no significant effect in mathematics. This difference in impact on student achievement between language arts and mathematics may suggest that there are a number of considerations that can be made here including the instrument that was used and its effectiveness to assess mathematics lessons, the nature of the post-observation conference, and the efficacy of the mathematical experience and ability of the evaluator. Is it possible that the use of a mathematics specific

instrument or the assessors possessing a strong mathematics background might be better able to make the post-observation sessions more meaningful and result in statistically significant increases in mathematics? The present study concentrated on one participant where focused feedback was given from someone with specialized knowledge in elementary mathematics. Three lessons were observed and feedback provided with the aid of a mathematics specific tool.

Marzano's Teacher Evaluation model (Marzano, 2011) provides another option for evaluating instruction. The instrument measures teachers on four domains: (1) classroom strategies and behaviors, (2) planning and preparation, (3) reflecting on teaching and collegiality, and (4) professionalism. The central idea supporting the Marzano evaluation tool is the belief that evaluation systems that foster learning will be different from those that measure teacher competence (Marzano, 2011). This suggests that different observation tools are able to provide varied types of information to the observer. For the Marzano tool, improved student performance is also a goal, since through the practices that are aligned with the tool, teacher expertise should be developed. The four domains are further divided into 60 elements that define a knowledge base for teaching and a framework that allows for the expertise of the teacher to be developed gradually (Learning Science International, 2016).

While using frameworks such as Danielson's and Marzano's can yield information to evaluators to assign grades and provide some information on readiness for certification among PSTs, the PSTs may be better served if there is an understanding of how these tools can help them improve their teaching (Roegman, Goodwin, Reed, & Scott-McLaughlin, 2015). Considering this, there are a number of classroom evaluation tools for mathematics that exist that are able to determine the nature of mathematics classroom instruction. Researchers are able to use the tools to focus on analysis of key features of intervention or treatment (Boston et al.,

2015). Such mathematics observation tools include The Reformed Teaching Observation Protocol (RTOP) (Sawada et al., 2002), The Mathematical Quality of Instruction (MQI) (Hill et al, 2008), and the Instructional Quality Assessment (IQA) toolkit (Boston et al., 2015).

The Reformed Teaching Observation Protocol (Sawada et al., 2002) is a mathematics observation tool in the form of a 25-item Likert scale questionnaire that examines instructions under three dimensions: Lesson Design and Implementation, Content, and Classroom Culture. The RTOP may be used by researchers to assess videotaped or live lessons. The RTOP has been used for more than 15 years in mathematics and science in a number of research studies. The more observations done, the more improved the statistical validity of the use of the instrument. Additionally, the greater the score, the better the indication of best teaching practices.

The Instructional Quality Assessment toolkit (Boston, 2012) is designed to measure the quality of mathematics instruction at scale. The IQA toolkit examines the instruction of mathematics through two constructs of Academic Rigor and Accountable Talk further sub-divided to more specific areas of a lesson that contains 4 scales per rubric. Training is recommended for use with the IQA toolkit where observers take detailed notes that are used to complete the rubric immediately following instruction.

The Mathematical Quality of Instruction (Hill et al, 2008) is an instrument designed to measure the nature of mathematical content available to students during instruction. The instrument is designed around five dimensions of Classroom Work Connected to Mathematics, Richness of the Mathematics, Working with Students and Mathematics, Errors and Imprecision and Common Core Aligned Student Practices. These dimensions are further divided into

subscales. To use the MQI coders must also be trained and it can be used to measure instruction regardless of the instructional approach used.

Boston, et al. (2006) asserted that while all the instruments have been validated for use in mathematics education research, specific attributes related to the instruments make them suited for unique settings and purposes. Although the RTOP was designed for mathematics and science application, the nature of the instrument is not content specific and may be applied to other content areas as well because it is more generally designed to evaluate reform-oriented practices. While MQI assesses rigor and richness in mathematics lessons, it is not designed for reform-oriented practices; however, it is able to be used to evaluate across a number of instructional approaches and is also specific to video assessment. The IQA toolkit draws the observer's attention to the presence of reform-oriented strategies in mathematics instruction such as cognitively-challenging tasks, task implementation, and discussion. The researchers also assert that the IQA toolkit may be ideal to evaluate PSTs at the elementary level. Although these three evaluation tools would be useful in research, for the purpose of this research on PSTs where instruction will be measured "at scale", which means that the instruction will be measured at the time of instruction, and where the tool is being used for its qualitative components in the feedback process, the IQA toolkit is the most suitable tool to be used.

The IQA Toolkit. Out of the need to have evaluation tools that present the opportunity to determine the important aspects when observing a lesson, the IQA toolkit was developed. Initial attempts to develop the IQA toolkit came out of the efforts by Matsumura (2000) to assess collection of homework and students' work. Active development, however, was done at the Learning Development and Research Center (LDRC) at the University of Pittsburg beginning in

2002 (Junker et al., 2006). Junker et al. went on to explain that the IQA toolkit was conceptualized around specific guidelines to capture instructional practice that integrates strong pedagogical knowledge and rigorous subject matter knowledge designed to measure the quality of mathematics instruction at scale.

The tool was originally developed for mathematics and reading comprehension at the elementary levels. The IQA toolkit was used to evaluate two lessons per teacher as well as assignments and written work measured on four indicators of cognitively-challenging tasks, task implementation, opportunities for students' explanation of thinking and reasoning, and teachers' expectations for student learning (Boston, 2012).

The Instructional Quality Assessment toolkit offers a way not just to assess teachers numerically, but also to provide explanations of the level of teaching based on observed characteristics. The toolkit was developed to capture both statistical and descriptive data about the nature of mathematics instruction and students' opportunities in the classroom (Boston, 2012) and measures instructional quality through lesson observations, assignment collection, and students' work (Boston et al., 2015). Two central constructs, or groups, to the IQA toolkit are Academic Rigor and Accountable Talk.

Firsthand account of the practices in the classroom, including what students and teachers are engaged in, as well as the process in which mathematics is taught and learned, can be captured using the IQA toolkit (Boston, 2012). The construct of Academic Rigor, seeks to capture the rigor of the instructional task, the rigor of the discussion, and the overall implementation of the task (Boston, 2015).

The task that students are provided, whether conceptual or procedural, provides opportunities for students to demonstrate their understanding of mathematics (Stein & Lane, 1996). However, teachers have experienced difficulty in maintaining the level of the task and it often decreases over the period of implementation (Stein, Remillard, & Smith, 2007). Maintaining the level of the task is of importance due to gains in student achievement being tied to it (Stein & Lane, 1996). Academic Rigor is measured through the rating of sub-constructs or sub-groups namely the Potential of the Task, Student Discussion After the Task, Questioning, and Implementation of the Task.

The Potential of the Task rubric identifies the highest level of thinking a task can potentially elicit from the students (Boston, 2012; Boston et al., 2015). The Potential of the task is rated by taking into consideration the highest level of cognitive thought needed to produce a comprehensive response to the proposed task (Boston, 2012). The Implementation of the Task rubric assesses the level of rigorous thinking in which students are engaged during solution of the task, focusing on the highest level in which students are engaged during the process of working on the task and discussing the task (Boston, 2012). Important to the Academic Rigor construct as well is the level of questioning that teachers use to engage students, exceeding trivial, low-level questions to asking questions that are academically relevant to elicit conceptual mathematical responses. Cognitively challenging tasks that are given to students can be maintained through the level or quality of questions students are asked (Boston, 2012). Central to the cognitively challenging task is the discussion following the introduction of the task.

Engaging students in mathematics discussion following the introduction of the task provides the opportunity for students to engage in reasoning and thinking at a higher level. The whole-group discussion provides the opportunity for students to advance their thinking and

reasoning skills in mathematics (Boston, 2012). Accountable Talk is captured by the sub-constructs of Teacher Linking, Student Linking, Teacher Press, and Student Providing. The Accountable Talk rubrics capture teachers' connections of mathematical ideas, students' connections of mathematical ideas, teachers eliciting responses from students through pressing for conceptual explanations, and students providing mathematically-sound responses. Linking describes mathematical talk in the classroom that connects ideas that are expressed. The Teacher Linking rubric captures the teacher's talk moves to extend, analyze, or critique mathematical contributions of others while the Student Linking rubric captures the efforts made by the students to make mathematical connections as well as comparisons to the work of others in the classroom (Boston, 2012). Teacher Press refers to the teacher moves to make students accountable for their contributions to the discussion by providing justifications to claims or providing extended responses. Student Providing assesses the extent to which students responded satisfactorily by justifying their claims or providing extended responses.

The Accountable Talk rubrics are underpinned by the concept of mathematical discourse, categorized by Imm and Stylianou (2012) as low discourse, high discourse, and a combination of low and high discourse. In high discourse classrooms, teachers value student ideas and facilitate rich mathematics conversations where students are included in purposeful discussions. Low discourse patterns focus on the talk moves dominated by teachers with minimal effort from students. The researchers describe the combination of high and low discourse as hybrid discourse patterns where discourse patterns vary during instruction.

Table 2 provides a synopsis of a few groundbreaking studies that have been conducted using the IQA toolkit. It highlights, as well, that the element of feedback was not included in

these studies. Studies highlighted in the table are also quantitative, making inferences based on numerical values obtained.

Table 2: Examples of Research Studies that used the IQA toolkit

Author(s)	Year & Methodology	Title	Central Claim	Methods	Results/conclusion	Possible Gaps
Janet C. Quint, Theresa M Akey Shelley Rappaport Cynthia J. Willner	2007 Quantitative	Instructional leadership, teaching quality, and student achievement: Suggestive evidence from three urban school districts	Providing instruction-related professional development to principals create chain of events to improve teaching and learning in their schools	49 elementary school principals provided with PD to pass on to teachers	Involvement of principals in PD is not associated with improved scores for math but for English as measured by the IQA toolkit.	- MCK for principals is insufficient - No feedback element - Not qualitative - Not PSTs
Melissa D. Boston and Margaret S. Smith	2009 Quantitative	Transforming Secondary Mathematics Teaching: Increasing the Cognitive Demands of Instructional Tasks Used in Teachers' Classrooms	PD training improves teachers' selection of cognitively demanding tasks	18 secondary math teachers, participate in PD- lesson observation and interviews	High level tasks were selected more frequently after participation in Professional development	- Not elementary - No feedback element - Not PSTs - Not qualitative
Lindsay Clare Matsumura Helen E. Garnier Sharon Cadman Slater Melissa D. Boston	2008 G-Study (Quantitative)	Toward Measuring Instructional Interactions "At-Scale"		34 Grade 6 and 7 teachers observed over 2 consecutive days	- As few as 2 observations might yield a reliable estimate (same rater on both occasions)	- Not elementary - No feedback element - Not qualitative - Not PSTs
Melissa D. Boston Anne Garrison Wilhelm	2015 Quantitative	Middle School Mathematics Instruction in Instructionally Focused Urban Districts	Cognitively challenging tasks decrease during implementation	114 middle school mathematics classrooms	Ambitious math instruction focused on students learning math with understanding can decrease the achievement gap - teachers need support to (a) maintain students' opportunities for thinking, reasoning, and problem solving throughout lesson implementation	- Not elementary - No feedback element - Not qualitative
Brian Junker Yanna Weisberg, Lindsay Claire Matsumura Amy Crosson, Mikyung Kim Wolf Allison Levison, Lauren Resnick	2006	Overview of the Instructional Quality Assessment	The IQA toolkit is able to capture various aspects of classroom instruction tied to instructional quality	16 mathematics lessons and 14 reading lessons observed on the construct of Academic Rigor	The IQA toolkit is able to determine features of instruction that have changed due to professional development	- No feedback element - Not qualitative

Table 2 also highlights some of the gaps that this current research aims to fill. What is evident in these studies is the lack of research using the IQA toolkit at the elementary level as most of the studies reviewed focused on middle and high school classrooms. Important to note is that there is a missing element of feedback in the research designs and the analyses were quantitative in nature.

While studies have highlighted that the IQA toolkit can be used to measure the presence of ambitious mathematics instruction, they did not capture how it could be used as a tool to support the feedback process. This research seeks to add to the literature on the IQA toolkit to possibly further highlight the tool as one that is able to capture the necessary moves of the PST and the students in the mathematics classroom in addition to being used in the feedback process. This addition to the research around the IQA toolkit is crucial when considering that feedback is a key element in the internship process (O'Shea, Hoover, & Carroll, 1988; Benson & Hooton, 2014).

Theoretical Framework

Feedback intervention, as defined by Kluger and DeNisi (1996), refers to the "actions taken by some external agent to provide information regarding some aspect of one's task performance" (p. 255). Prior researchers of feedback intervention have not been able to find conclusive results related to feedback intervention. The disparity in research results has led many researchers to assume that feedback interventions consistently improve performance (Kluger & DeNisi, 1996). Kluger and DeNisi stated that the absence of a theory related to feedback contributed to the assumption that feedback results in improved performance at all times, which may not be true and therefore they proceeded to clarify research and develop what

is called the Feedback Intervention Theory (FIT). They explain that FIT represents a wide range of feedback intervention (FI) to include areas such as test performance, memory tasks, and physical tasks. Their definition, however, does not include natural feedback processes that operate without external intervention, task generated feedback that is obtained without intervention, personal feedback that does not relate to task performance, and self-initiated feedback-seeking behavior. In essence feedback must come from an external source to be considered in the theory. An example of the natural feedback process as explained by the researchers includes feedback on a loop control system that gives a feedback signal for example in an electrical system. An example of task generated feedback without intervention could be a farmer who turns a water system on for too long and realizes the field is flooded. The farmer is able to correct the behavior without external intervention. Personal feedback refers to the focus on the personality of the individual as the basis for the feedback rather than feedback on the task that the person has performed. Feedback-seeking behavior refers to a conscious effort by someone to determine the adequacy of his behavior as it relates to attaining a goal (Ashford & Cummings, 1983 as cited in Crommelinck & Anseel, 2013). Someone who requires to know how he performs from another as it relates to attaining a goal is demonstrating feedback-seeking behavior that is initiated by him, therefore called self-initiated feedback seeking behavior.

Kluger and DeNisi stated that control theory (Carver & Scheier, 1981) and goal setting theory (Locke, Shaw, Saari, & Latham, 1981) contributed significantly to their development of FIT as they include those components of feedback in their theory. Control Theory (CT) is a general approach to understanding self-regulating systems and provides a model of self-regulation that is useful in the analysis of human behavior. Carver and Scheier propose that the control process is represented by a feedback process. This feedback process called the feedback

loop focuses on negative feedback because its function serves to negate or reduce deviations from a comparison value. The input senses the current condition when compared against a discrepancy, or deviation, from what is expected. If a discrepancy is apparent and there is behavior in response to the discrepancy, the behavior performed is the output function, the goal of which is to reduce the discrepancy. For example, if a teacher is teaching a concept in mathematics in which students are expected to respond conceptually, if the students are able to provide a conceptual response, that response is considered the "expected state". If a student provides a response that is not conceptual, this is a deviation from the expected state. The teacher immediately responds in an effort to elicit the conceptual response, which is the output function. The teacher may respond by pressing for conceptual responses by asking academically relevant questions, in an effort to reduce the discrepancy. For teachers who are more experienced, the process to do this likely takes less effort so the discrepancies are kept at a minimum. Teachers who are less experienced would likely have more difficulty getting students to answer conceptually.

The work of Locke and colleagues in developing the goal setting theory spans over several decades (Locke, Shaw, Saari, & Latham, 1981; Latham & Locke, 2002; Locke & Latham, 2007). They explain that a goal is something that an individual is trying to accomplish. This may also be described as a performance standard, task, or objective. In many goal-setting studies, the term goal refers to attaining a certain standard of proficiency on a task. Important to note according to Locke and colleagues is that goals that are more difficult to achieve would be met with greater effort to accomplish them than those that are not. Goal theory predicts that harder goals lead to better performance than easy goals, despite the lower probability of those goals being fully attainable (Locke, Shaw, Saari, & Latham, 1981). Although the control and

goal setting theories were integral in the development of FIT, other theories that contributed also had feedback as a component.

Kluger and DeNisi (1996) proposed what is termed as a hybrid theory of goal setting, control, and other theories called FIT that combines aspects of these existing theories and uses processes that are not addressed in the existing theories. The Feedback Intervention Theory as stated by Kluger and DeNisi is based on five arguments: "(a) behavior is regulated by comparisons of feedback to goals and standards, (b) goals and standards are arranged hierarchically, (c) attention is limited and therefore only feedback-standard gaps that receive attention actively participate in behavior regulation, and (d) attention is normally directed to a moderate level of the hierarchy and FIs change the locus of attention and therefore affect behavior" (p. 259). Hierarchy as used here surrounds the idea that there is a range from task at the lower level and self at the higher level and that feedback results in attention being placed more on the task or on self. Kluger and DeNisi alluded to the idea that the effectiveness of feedback intervention decreases when attention is focused on self rather than what should be accomplished.

Kluger and DeNisi (1996) put forward the argument that a basic mechanism in behavior regulation is the response to a feedback-standard comparison alluded to in both the goal setting and control theories. Both theories view behavior as goal directed since to achieve goals and standards people use feedback whether or not it was provided by an intervention. This feedback is used to evaluate the extent to which the goal is achieved. The result of a comparison of feedback to a goal or standard is followed by a feedback sign relative to the goal, whether positive or negative. When there is a discrepancy and feedback is provided, the response falls in

the categories of the two theories of goal setting and control theory. If associated with control theory, there are a number of options that may be considered. The discrepancy could be eliminated resulting from a behavior change so that future feedback will also be changed. Additionally, changing the standard so it reflects the current feedback or omitting the situation from which the discrepancy arises may also be considered as an option.

For the goal setting theory, Kluger and DeNisi (1996) report that the options are different as the individual can try to attain the goal, change the goal, reject the feedback, or abandon the commitment to the goal. The response is usually that people try to eliminate the feedback-standard discrepancy by attempting to meet the standard. This usually results in an increase in the rating of the previous assessment. Feedback intervention that reflects that an individual's performance has exceeded that standard results in the effort related to that standard being reduced or maintained while feedback intervention that is negative is likely to lead the individual to exert more effort than one who has received a positive response. An example of the response being associated with the goal setting theory is related to creating opportunities for students to connect mathematical ideas by identifying similarities, differences, or connections in methods of solutions or results. These opportunities can be described as linking behaviors.

While the literature on evaluation in mathematics itself is extensive, the literature surrounding detailed feedback in mathematics is not. The studies involving the use of the IQA toolkit, although used as well to determine the instructional quality of mathematics lessons, did not include the feedback process to the teachers being evaluated. The current study provides insight into the feedback process, through the use of the IQA toolkit, in an effort to understand if there is any influence, and the nature of such to the instructional quality of a PST.

CHAPTER 3: METHODS

In this section of the dissertation, the sample is described. The process to collect data from the sample is also outlined and the steps employed in the analysis of the data are documented. Included in this chapter, as well, is the description of how the data were used to answer the research questions.

Research Design

The idea for the study came from a previous study with which the researcher was involved that measured the quality of instruction among practicing elementary teachers. The study, titled Replicating CGI in Diverse Environments (https://ies.ed.gov/funding/grantsearch/details.asp?ID=1295) used the IQA toolkit to measure the instructional quality of practicing elementary teachers using video recording. The study, however, did not focus on the element of providing feedback or other sorts of follow-up based on the use of the tool to teachers who were involved.

The study reported here used the critical element of feedback after the instructional quality had been rated using the IQA toolkit. This single case study sought to provide a description of the influence of feedback on the instructional practices in mathematics of a final year preservice elementary teacher (PST) from a large metropolitan public research university in a southeastern state in the Unites States of America (USA).

Research Questions

The research questions that were analyzed in this single case study are:

Main Question: How does feedback using the Instructional Quality Assessment (IQA) toolkit influence the mathematics instructional practices of a final year preservice elementary teacher?

Sub-questions

1. In what ways does feedback using the construct of Academic Rigor as measured by the Instructional Quality Assessment toolkit influence the mathematics instructional quality of a final year preservice elementary teacher?
2. In what ways does feedback using the construct of Accountable Talk as measured by the Instructional Quality Assessment toolkit influence the mathematics instructional quality of a final year preservice elementary teacher?

Population and Sampling

The participant for this convenient, purposeful, and homogeneous sample consisted of one PST in her final year internship in the spring 2017 semester. Additional selection criteria for the participant included the following:

1. A student who had successfully completed a course focused on elementary school mathematics.
2. A PST whose internship was in a school where solo teaching was allowed rather than an environment that exclusively called for co-teaching.

There was no need for random selection since the participant that was used identified that she wanted to be included in the study and she was one of two participants who volunteered. The other participant completed one of three cycles of data collection and decided not to continue and therefore the data for that participant were not included in this study. Initially the

researcher intended to include four participants but encountered difficulties recruiting participants for the study. This is likely due to some initial confusion regarding the request to participate in the research and to the stress experienced by PSTs during this stage of their teacher preparation process. The pressure of finding time to meet and also to complete reflections might have dissuaded further participation.

Data Collection Procedures

Prior to initiating the research, permission was sought and received from the Institutional Review Board (IRB) (Appendix B) of the university the researcher attends and from the school district within which the data were collected. The participant was recruited from the population of final-year preservice elementary teachers who were completing their second and final internship in the spring semester of 2017. The director in the department of clinical experiences in the college of education who has the responsibility for deploying interns assisted in the recruiting process and sent emails to potential participants. Principals of elementary schools were also informed of the permission granted for research to be conducted in their schools (Appendix C).

After encountering some difficulty in recruiting participants, two students stated that the request was unclear due to the word "tool" used in the request, which was interpreted as a manipulative, and were therefore dissuaded from participating in the research. Potential participants also stated that their workloads were too high for them to take on an additional task to engage in the research. Two students began the research and one of the two determined that her work load was too high to continue. Therefore since only one participant completed the entire data-collection process, only this participant's data were considered. Prior to the

collection of data, the researcher met with the participant to discuss the process, expectations, and participation requirements.

Instrumentation and Data Gathering

Data were collected through two primary means to ensure that data triangulation could aid in confirming the results (Creswell, 2013; Gall, Gall, & Borg, 2007; Hancock & Algozzine, 2006). Data were collected from the participant using the following data collection instruments and methods:

1. Lesson observation notes taken during the lesson, stating occurrences in the lesson delivery process.

   The notes recorded during the lesson aimed to capture the instructional moves by the PST, involving the extent to which opportunities arose for students to demonstrate conceptual understanding, and students' responses to these situations. Essentially, the occurrences in the classroom that were likely to be reflected in the IQA toolkit were recorded.

2. The IQA toolkit scoring rubrics (See Appendix A).

   The IQA toolkit rubrics were used to rate the participant each time she was observed. While the individual scores were considered, the qualitative description corresponding with the score was the more important consideration since the key feature of the feedback was based upon these data. The detailed descriptions at each score level provided data favorable for qualitative interpretations of the IQA toolkit (Boston et al., 2015).

3. Written reflections by the participant.

The written reflections by the participant were completed after each lesson using prompts related to lesson delivery, the PST's perceived performance, and engagement of students in the class through a feedback protocol (Appendix D).

Audio recording and transcription of feedback process. The feedback process was audio recorded and partial transcriptions completed as needed to determine aspects of the feedback that were important to answer the research questions. All sections of the feedback sessions were not transcribed because there were instances in the discussion that did not directly provide data for the study.

Steps for Data Collection Cycles.

Usable data were collected from a single participant who went through three data collection cycles reflected in the diagram in figure 2. The number of observations was consistent with suggestions by Wilhelm and Kim (2015), where the teacher had whole class discussion after the task. Wilhelm and Kim suggested that at least three observations per teacher were needed to reliably measure a teacher's instruction using the IQA toolkit. Also reported was that more observations are needed for classes that do not have whole class discussion following the task.

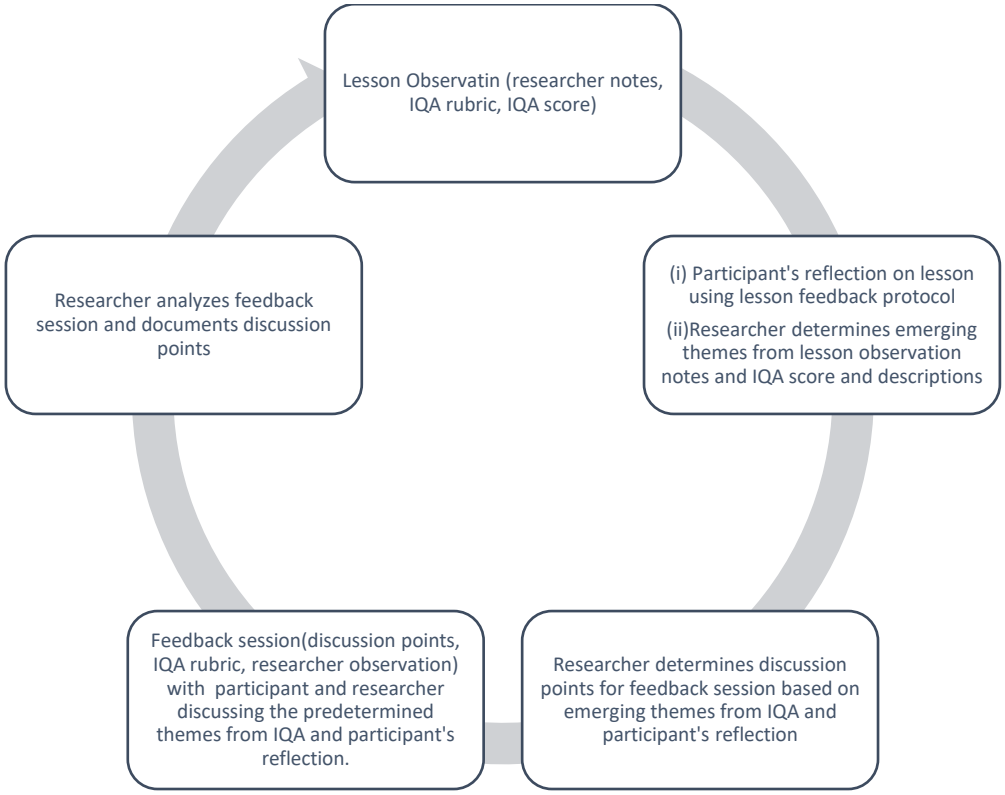


Image by author

Figure 2: The Data Collection Cycle

Each cycle took place over a number of days that varied based on the availability of the participant among other factors. However, each cycle was fully completed before another started.

Step 1: The PST taught a lesson observed by the researcher. The researcher took copious notes during the lesson documenting the conversations during the whole class discussion. Although there were instances when the teacher would talk to students in small groups, these were not recorded because the IQA toolkit was designed to measure discussions that occur during the whole class.

The notes that were recorded were used to determine what aspects of the lesson contributed to the rating of the IQA toolkit rubrics. All that was considered talk was documented. Talk would include each question, answer, statement, explanation, and justification which was taken as researcher notes.

Step 2: The PST wrote a reflection of the lesson without specific direction from the researcher. However, the PST was provided with a reflection protocol (Appendix D) as a guide. The protocol prompted reflection on the achievement of the objectives, if there was deviation from the lesson, and how the deviation was addressed. There were also prompts to focus the participant's reflections on the task and if it was a rich task, how it was implemented during the lesson, and if the way it was implemented was effective. The PST was also invited to reflect on whole class discussion and its effectiveness and to state if she asked academically-relevant questions.

Step 3: The researcher used coding themes (later described) to analyze the reflection of the PST. The themes that emerged were noted and those themes as well as the corresponding qualitative description of the IQA toolkit scores were used to determine the focus of the feedback process. The data were concerned with mainly words, from the researcher's notes and the PST's reflection. The researcher had to read through the observation notes and compare them with the qualitative descriptions from the IQA toolkit to assign a score. This process was also used to determine the areas that need to be discussed in the feedback session. A theme would be considered when a sub-construct was not given the highest rating.

Step 4: The feedback session was video recorded. The PST chose the venue to do the reflection; the principal offered a conference room as one option. The entire discussion was

recorded to capture the discussions that took place for the feedback session. Aspects of the feedback session were transcribed that could be usable data. Conversations in the feedback session that were unrelated to the IQA toolkit were not transcribed. The contents of the feedback sessions were eventually compared for any notable differences in how these sessions were conducted.

Step 5: The feedback session discussion points were noted and portions of the video recording of the feedback session pertinent to the research were transcribed.

Data Analysis Methods

The data analysis was ongoing during each phase of the research. The participant's data were analyzed as described in table 3. The table provides an overview of the analysis that was ongoing throughout the data collection process.

Table 3: Data Analysis Procedures

Data	Data analysis procedures
Researcher observation notes	The researcher documented the discourse that took place during instruction. Notes corresponding to each sub-construct were identified and compared to the description on the IQA toolkit.
Participant's written reflection (after lesson)	The occurrences during instruction that the participant reflected on were documented, specifically, the occurrences related to elements of the IQA toolkit. The important ideas that were brought out in the reflection as prompted by the reflection protocol were documented. The responses were coded in themes, based on the reflection of the participant. There was also specific attention given to the language of the reflection to determine if there were any words or phrases specific to the IQA toolkit that were used in her reflection. The references made to aspects of Academic Rigor and Accountable Talk were documented and placed under sub-constructs in order to determine the specific areas related to the IQA toolkit that the participant included in her reflection. For example, if the participant reflected on her questioning effectiveness, this would be related to the Academic Rigor construct, and specifically to the sub-construct of questioning.

Video recording of feedback session	The feedback sessions were transcribed as needed and the discussion was coded for themes. Transcripts from the feedback session were determined by the points of the discussion that were directly related to the sub-constructs of the IQA toolkit.

Application of Theoretical Framework

As stated by Kluger and DeNisi (1996), feedback intervention refers to the actions taken by some external agent to provide information regarding one's performance. In this research, the actions were taken by the researcher assuming the role of the external agent who responded to the actions of the participant through observation and written comments. The FIT framework was used in the analysis phase to determine the responses of the participant to the feedback. Questions from the theory that were answered included the nature of the feedback that was given, whether positive or negative, and the resulting behaviors. For example, if the Potential of the Task was being considered, the discrepancy would surround the nature of the task. If the task was rated at a four (indicating a high cognitive demand task), the feedback would likely be positive. Considering the framework, the participant then had the opportunity to focus on other goals or aspects of instruction. The two choices that were likely to be made were (1) other goals were set to be achieved or (2) there was no need to set other goals and effort of the participant in this area was reduced. If the feedback was negative, it was assumed that there would be an increase in effort from the participant. Two things would have to be determined at that time (1) if the increased effort reduced the discrepancy and what was the result or (2) what happened if the increased effort did not reduce the discrepancy? Did the participant shift focus to another area and was that shift related to self or related to another area such as other aspects of student learning. If the PST responded positively, then that would be related to the goal setting theory since the PST was trying to eliminate the discrepancy. If the discrepancy was in fact eliminated, the link to the control theory was more apparent since one of the responses to the control theory is to eliminate the discrepancy completely.

If there was positive feedback, an opportunity was presented to attain other goals. An observation then could be made to determine if the standard was raised, meaning that there was more to be achieved or if the teacher felt that the effort did not need any additional attention. Of those two options what were the results? Was there an increase in effort or was the level of the effort reduced? If there was negative feedback the opportunity was provided to increase the effort. The observer would then determine if the increased effort reduced the discrepancy or if the teacher shifted focus from that discrepancy.

It therefore meant that each follow-up lesson observed was compared to those observed previously to make a decision regarding the response to the discrepancy. This determined the possible changes in the behaviors of the participant and whether she was influenced by the use of the IQA toolkit. It was then determined if the claims that were made were supported by the qualitative descriptions of the IQA toolkit and the researcher's observation notes. For example, the participant is observed and writes her reflection on the instructional process and feels she has provided rich tasks, which when compared to the observation notes from the researcher and the subsequent rating from the IQA toolkit, does not reflect the same thought as the participant. The qualitative descriptions that corresponds with a score of four would then be discussed and the PST would try to rewrite a problem that could be given to the students to reflect a four on the IQA toolkit rating. Additionally, the aspects that make the task rich would be reviewed.

CHAPTER 4: DATA ANALYSIS

The purpose of this study was to determine if feedback influenced mathematics instruction, specifically through the use of the Instructional Quality Assessment (IQA) toolkit. Using the approach of a single case design, two data sources were used repeatedly to determine themes that arose from analysis of these sources. The researcher sought, per her research questions, to determine if feedback influenced the mathematics instructional quality of a preservice elementary teacher on the constructs of Academic Rigor and Accountable Talk as defined by the IQA toolkit. The themes, therefore, that arose were linked to the rubrics in the IQA toolkit to determine if there were any influences that could be observed in the quantitative rating as well as the qualitative descriptions on the evaluation instrument. Influences could be determined by changes in score over feedback cycles which could also be linked to the changes in the description of what was observed in the class as explained by the qualitative aspects of the IQA toolkit.

Data analysis was therefore conducted throughout the data collection process as the results from each analysis cycle from the three data sources influenced the nature of the feedback that was given to the participant. The content of the feedback was determined by the themes that arose from the analysis of the notes from the observation of the participant's teaching and the analysis of her reflection. Since the IQA toolkit was significant to the research, the themes related to the sub-constructs were considered throughout the analysis process. Therefore the reflection of the participant was consistently linked to the IQA toolkit rubric to determine if there were connections to any of the sub-constructs of the rubric. For example, the participant in her first lesson did not show any evidence of teacher linking other than revoicing students'

contributions and she did not have comments that could be connected to Teacher Linking in her reflection. Since there was no evidence of Teacher Linking from the reflection, the theme then emerged from the IQA toolkit rubric. There was extended discussion in the feedback session related to Teacher Linking. Questioning is a part of the IQA toolkit rubric and this theme emerged from the reflection of the participant and the IQA toolkit rubric result. Therefore in this instance the theme emerged from the two sources. The absence of identifiers of a construct as well as what was identified by data collection contributed to emerging themes.

After the feedback session, the participant began to prepare for another cycle teaching; another lesson that was observed and rated using the IQA toolkit. She also wrote a reflection on her new lesson. After the analysis was completed for the new lesson and the reflection, feedback was given and a third cycle of data collection began. The participant was involved in three data collection cycles for the duration of the research. The data-collection process was conducted over five weeks.

Description of the Participant

For the purpose of this research, the case study teacher will be identified as Miss S. Miss S was approximately 21 years of age and was in her final year at a large, urban university in the southeastern United States, working towards her Bachelor of Science in Elementary Education. Having completed four years of studies (2 years outside the university), Miss S was enthusiastic to start her teaching career and believed that she would be able to help students learn, using strategies which focus on them being active learners.

While at her university, Miss S completed at least 2 courses in elementary mathematics. She completed at least one content and one methods course. On meeting Miss S for the first time the researcher asked about the mathematics classes she had taken. She had difficulty remembering the content course but she remembered when the researcher mentioned base eight. She laughed when she remembered. The course being mentioned is an elementary mathematics content course that exposes students to mathematics concepts, including whole number operations using base eight, focused on students creating meaning and understanding.

In one of the feedback sessions Miss S mentioned that she loved mathematics and also liked teaching the subject and would want to teach her students in ways that would develop their love for the subject as well. Miss S was eager to learn anything that would improve her mathematical knowledge. She said she was very excited to teach using current teaching practices, although she enjoyed the job she held in a restaurant. Miss S said she had never been involved in any formal professional development outside of university requirements. She did not hesitate to participate in the study and was willing and excited to learn. At the time of the study, Miss S had the remainder of her internship to fully complete her studies, and had successfully completed her state certification examinations. Miss S was the ideal participant for the study because she was in her final year in her course of study and was on her final internship before being certified as a teacher.

The Classroom

From the classroom door, the teacher's desk faced the door, the whiteboard was to the right, and the students' desks were to the left. There were student activities in the form of games, puzzles, books, and manipulatives around the class for different subject areas.

Manipulatives such as rulers, snap cubes, and play money were available for use during and outside of mathematics instruction time. The students' desks were arranged in groups of four, which allowed for students to share manipulatives. During the lesson, students were able to converse with each other as instructed by Miss S.

When I met Miss S, she had already completed four weeks of internship. Miss S was required to teach all subjects to all of the students in her class except for those who were removed for special consultations with the intervention specialists, some taking place during her teaching times for mathematics. Her class had between 14 and 17 students depending on whether students were removed for the specialized sessions or not. Miss S was assigned to the second grade with a teacher who was previously involved in professional development in Cognitively Guided Instruction (CGI). While it is not clear the other types of professional development the supervising teacher may have done, CGI is a professional development program that resulted from research into students' thinking as a guide to designing instruction (Carpenter, Fennema, Franke, Levi, & Empson, 2000). The observable norms of the class at the start of the researcher's observation of Miss S's instruction included students contributing to the lesson but most of the discussion was dominated by Miss S, students worked with partners or in larger groups, and students tended to try to use their own methods to solve problems.

Analysis during Data Collection

Assigning Scores

Data were analyzed as soon as they were collected so that the initial analysis could inform the next steps in the data collection process, specifically in the feedback portion of the

cycle. After the lesson observation, the observation notes were used as evidence to assign scores for the IQA toolkit. Assignment of scores in each of the sub-constructs of Academic Rigor (Appendix A) was based on the opportunities for students to engage in rigorous thinking on content that presented some level of challenge measured through the task, its implementation, whole class discussion after the task, and student questioning. The sub-constructs of Academic Rigor have scales from zero to four. A score of zero indicates that there were no observable data and a score of four indicates that the level of rigor was high. Scores of one, two, and three are placed on the continuum indicating varying levels of rigor. The score levels of the IQA toolkit provide very detailed qualitative descriptions so that the rater is able to accurately score, as much as is possible, based on the descriptions provided, when compared to observations made of instruction. Accountable Talk measured through the sub-constructs of linking (teacher and student) and press (teacher and student), focused on whole class discussions that allowed students to present and justify their ideas, methods, and thinking. Similarly to Academic Rigor, Accountable Talk (Appendix A) has scales from zero to four, with zero indicating that the behavior was not observed and four indicating the minimum number of times an occurrence can be observed for the sub-construct to be considered high level. A score of one or two would be considered low level and a score of three high level, but missing some of the required mathematical connections to be considered a score of four. Scores were assigned based on the observed occurrences that contributed to each sub-construct. After scores were assigned, the researcher reviewed the occurrences, or lack thereof, that contributed to the assigned scores. These occurrences were coded under themes that followed the names of the sub-constructs such as “task potential” or “questioning”.

Feedback cycle 1

The following is an explanation of the first feedback cycle. First the second grade lesson is described as enacted, then a description of how the IQA toolkit scores were assigned is provided. The emerging themes of interest based on the lesson observation and then the reflection are shared. Finally, the contents of the first feedback session are outlined.

The Lesson and IQA Toolkit Results

The second grade class consisted of 17 students. However, three students were removed during mathematics sessions for special meetings with two intervention specialists who teach mathematics concepts in a small group setting outside of the classroom. Therefore there were 14 students regularly in mathematics lessons. During the lesson the supervising teacher was present but did not interrupt Miss S during instruction. Figure 3 demonstrates the basic setup of the classroom at the beginning of mathematics lessons. The +1 at the tables represents the students who were removed for intervention.

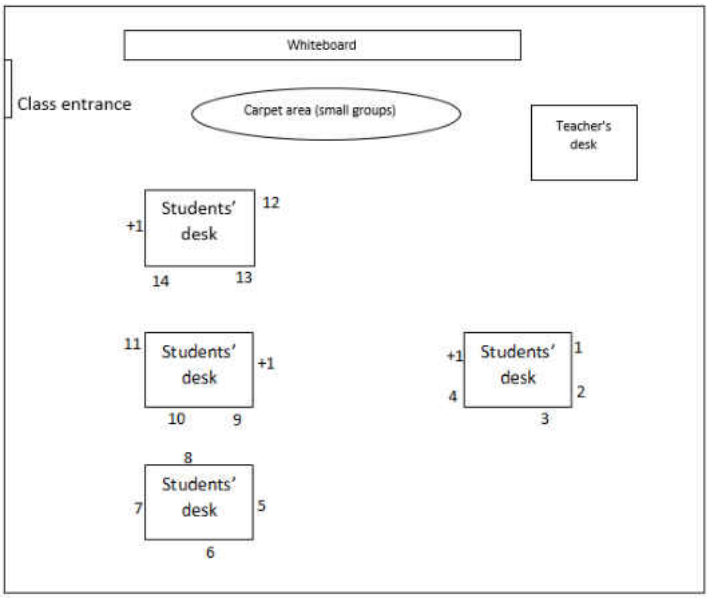


Image by author

Figure 3: Basic layout of the classroom

The teacher began the lesson by asking students if they remembered how to count by five. The class then skip counted by fives up to 35 and then counted by ones from 10-19. They counted by fives, tens, and ones several times starting from different numbers each time. This led Miss S into the task which required students to determine the sum of money they have in total given different combination of bills. The task for the lesson was presented as a set of similarly represented problems where students were asked to use counting methods of their choice to determine money totals given. First they were presented with the representation in figure 4, projected on the board, and asked to find the total but were not required to explain their thinking. Students worked independently at their desks, drawing representations of the bills they were given.

Image by author

Figure 4: Example of money for student task

All students were provided with the same activity regardless of perceived ability levels. The teacher circulated while the students worked on the problems. Students were asked to find money totals that used combinations of $20, $10, $5, and $1 bills. The activity lasted approximately 30 minutes with a total of four problems related to the task. After students had time to complete each problem, Miss S brought the class together to discuss the approaches of the students. The teacher, however, dominated the talk in the class although students were asked to share their strategies on one occasion, with occasional contributions requested from students.

The IQA Toolkit Rubric Results

The Academic Rigor Rubrics

Table 4: Lesson 1 Academic Rigor results

Sub-construct	IQA rubric rating
Potential of the Task	3
Implementation of the Task	3
Student Discussion after the Task	2
Questioning	1

Potential of the Task

The task was deemed appropriate for the grade level of the students as it related to the content, that it progressed in difficulty, and also that it required students to use connections from previous knowledge to solve the task. Students were given the opportunity to display solution methods that they chose. The teacher asked students to write down what they would do to answer the question but did not prompt students to provide evidence of their thinking. Evidence of students' thinking is demonstrated in their explanations and justifications of their methods and solution. Figure 5 and figure 6 show ratings for a score of three and a score of four respectively for Potential of the Task. The rating of three lists possible reasons the task would not receive a four even though the task was considered a good task for the level of the students. The rating of a four explicitly states that the students must be prompted for reasoning and understanding. The task received a score of three because the task itself lacked the critical element of asking students to display their thinking, that is, explicitly prompting for students' reasoning and understanding through explanation and justification (Boston, 2012). In the task given, students were provided with the bills that were projected on the white board. The solution path to determine the total amount of money was the decision of the student, but students were not explicitly asked to explain their thinking in their response.

Rating of 4

The task has the potential to engage students in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships, such as:

- Doing mathematics: using complex and non-algorithmic thinking (i.e., there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example); OR
- Procedures with connections: applying a broad general procedure that remains closely connected to mathematical concepts.

The task must explicitly prompt for evidence of students' reasoning and understanding. For example, the task **MAY** require students to:

- solve a genuine, challenging problem for which students' reasoning is evident in their work on the task;
- develop an explanation for why formulas or procedures work;
- identify patterns and form and justify generalizations based on these patterns;
-
- make conjectures and support conclusions with mathematical evidence;
- make explicit connections between representations, strategies, or mathematical concepts and procedures.
- follow a prescribed procedure in order to explain/illustrate a mathematical concept, process, or relationship.

Boston (2012). Used with permission of the author

Figure 5: Potential of the Task rubric rating of 4

Rating of 3

The task has the potential to engage students in complex thinking or in creating meaning for mathematical concepts, procedures, and/or relationships. However, the task does not warrant a "4" because:

- *the task does not explicitly prompt for evidence of students' reasoning and understanding.*
- students may be asked to engage in doing mathematics or procedures with connections, but the underlying mathematics in the task is not appropriate for the specific group of students (i.e., too easy or too hard to promote engagement with high-level cognitive demands);
- students may need to identify patterns but are not pressed for generalizations or justification;
- students may be asked to use multiple strategies or representations but the task does not explicitly prompt students to develop connections between them; students may be asked to make conjectures but are not asked to provide mathematical evidence or explanations to support conclusions

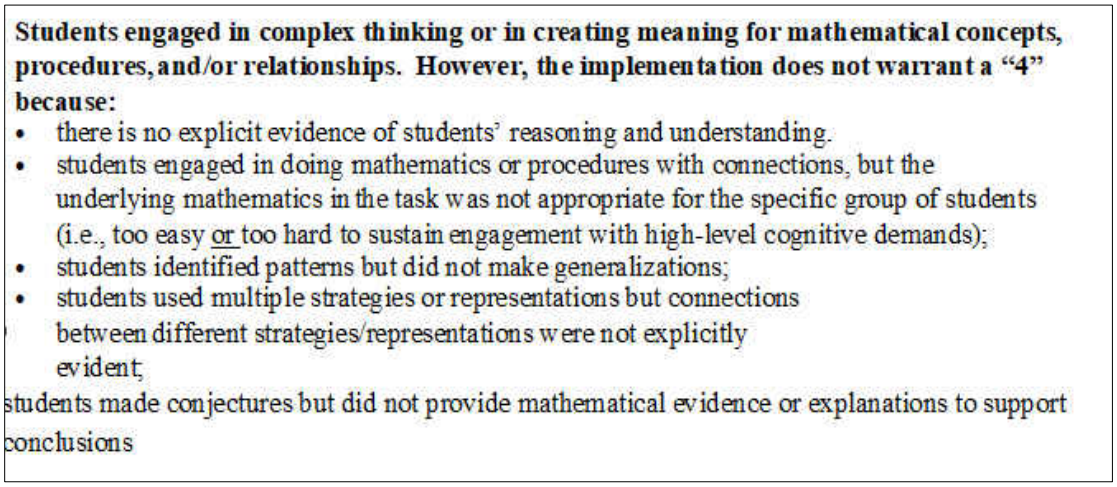
Boston (2012). Used with permission of the author

Figure 6: Potential of the Task rubric rating of 3

Implementation of the task

The implementation of the task refers to the opportunities that are provided to students to enable them to engage in higher order thinking and reasoning throughout the instructional process (Stigler & Hiebert, 2004). In considering the implementation of the task, the level of thinking required from students throughout interaction with the task was observed. The observer needed to determine if students interacted well with the task, if the students were having difficulties, and how Miss S, in this case, responded to such behaviors. The researcher had to consider if assistance given to students was too much so that it lowered the demand of the task, therefore making it easier. If the demand of the task is lowered opportunities for students to engage in productive struggle is reduced.

For the first lesson observed, Miss S received a score of three. Figure 7 shows the ratings corresponding to a score of three. The researcher concluded that the students were engaged in complex thinking, since they were required to make a choice regarding the method they would use on their own and they were allowed to find the total for the money combinations they were given by counting in different ways. However, Miss S did not encourage students to make connections with the mathematics discussed. There was an effort to support every student to have an opportunity to be involved in the lesson but the depth of students' understanding of the mathematics was not evident.



Students engaged in complex thinking or in creating meaning for mathematical concepts, procedures, and/or relationships. However, the implementation does not warrant a "4" because:

- there is no explicit evidence of students' reasoning and understanding.
- students engaged in doing mathematics or procedures with connections, but the underlying mathematics in the task was not appropriate for the specific group of students (i.e., too easy or too hard to sustain engagement with high-level cognitive demands);
- students identified patterns but did not make generalizations;
- students used multiple strategies or representations but connections between different strategies/representations were not explicitly evident;

students made conjectures but did not provide mathematical evidence or explanations to support conclusions

Boston (2012). Used with permission of the author

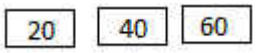
Figure 7: Implementation of the Task rubric rating of 3

There were times when there was an effort in making connections but it was Miss S who was making the connections for the students. In the excerpt below, Miss S asked students to determine the sum of two numbers. The students answered together and Miss S revoiced their

response. Revoicing (O'Conner & Michaels, 1996) is simply a repeat from Miss S of what the students contributed to the discussion.

Miss S: what is 40 plus 20 more

Students answer together: 60

Miss S: 60 (Teacher Revoice (TR)) because it's 40 plus two more tens so that's 40, 50, 60 and we would do  *(written on the board, representing increase in money values)*

Rather than asking students what counting method they used to arrive at 60, Miss S presented a method herself. This was observed at times during the lesson and therefore students' opportunities to respond were limited. The lack of students' responses potentially limited the opportunities for students to create connections with the responses of other students.

Student Discussion after the Task

Student discussion following the task focuses on the extent to which students show their work and explain and justify their thinking about important mathematical content (Boston, 2012). The Student Discussion rubric, then, captures talk within the whole group after the task has been completed. For the Student Discussion sub-construct Miss S received a score of two because, per the rubric (see Appendix A), students explained their work for solving the task, however, the rich discussion of why the strategies worked was not present. The discussion took the form of Miss S leading much of the conversation, although students explained their work in instances. Much of the students' responses were often one-word responses with no justifications

required and strategies were minimally discussed. If the student was correct, Miss S would go ahead and explain the justification to the rest of the students.

Questioning

The questioning rubric captures the level of academically-relevant questions asked of students through the whole class discussion. For lesson one, Miss S received a score of one (Figure 8). Questioning in the whole class discussion was limited to one-word responses from students. The teacher asked all students at least one question during the class but these questions were at times the same as the questions she asked everyone else in the class so the students' responses could have been simply repeating answers provided by others without them understanding the meaning. Figure 8 shows the explanation for a rating of one where procedural or factual questions are asked and are likely to produce one-word responses. For this problem students were given money combinations of $10, $5 $1, and $1.

Score of 1
The teacher asks procedural or factual questions that elicit mathematical facts or procedure or require brief, single word responses.

Boston (2012). Used with permission of the author

Figure 8: Questioning rubric rating of 1

The questions also did not require students to exhibit any higher order thinking. The following excerpt gives an example of one such occurrence. In this excerpt, three students were called upon to respond.

Excerpt

*Miss S: Thumbs up if you think you can count the next one. Israel** (all names are pseudonyms)

Israel responds: 10,15,16,17

Miss S: Point on your paper where you see the value and circle it. That's how you know how much each is worth.

Miss S: Who else wants to count? Sammy?

Sammy: 10,15,16,17

Miss S: Grey

Grey: 10,15,16,17

In the excerpt, three students were asked to respond to the questions. The teacher moved on from one response to another without any discussion or further requirements from the students. For the questioning rubric to reflect a higher score, the teacher would need to ask questions that require students to demonstrate higher order thinking.

The Accountable Talk Rubrics

Table 5: Lesson 1 Accountable Talk rubric results

Sub-construct	IQA rubric rating
Teacher Press	1
Student Providing	1
Teacher Linking	1
Student Linking	1

The Accountable Talk rubrics of Teacher Linking, Student Linking, Teacher Press, and Student Providing describe information on the quality of the discourse that occurred in the classroom in response to work on the task (Michaels, O'Connor, & Resnick, 2008). The quality of the discourse for this lesson was measured during the whole class discussion following the task. The qualitative descriptions of the sub-constructs were used as a guide to assign ratings for the Accountable Talk rubrics.

Teacher Linking

Teacher linking captured opportunities set up by Miss S for students to connect each other's mathematical ideas. Mathematical ideas set up by Miss S that allowed students to critique or analyze the contributions of others were captured under Teacher Linking as well. The teacher received a score of one for this sub-construct because efforts to link were evidenced only by revoicing the students' contributions to the discussion. For lesson one, Miss S revoiced students' responses at least five times during the lesson. A score of two for this sub-construct means that at least once during the lesson there was evidence from Miss S of an opportunity to link students' ideas but the effort was not followed by Miss S actually facilitating the link.

Students' Linking

Students' linking refers to attempts made by the students to connect mathematical ideas to their peers' or to Miss S's ideas. This is evidenced in revoicing a peer's contribution or more deliberately agreeing or disagreeing with peers based on mathematical ideas presented. Students are also expected to provide justification for their position on the contribution of their peers. For

lesson one, the students did not make any connections to each other's mathematical work and therefore no evidence of Student Linking was observed and therefore the score for students' linking was one. This sub-construct did not have a rating of zero because zero means there was no mathematical discussion, according to the rubric in the IQA toolkit.

Teacher Press and Student Providing

Teacher Press refers to the efforts by Miss S to elicit further explanations, clarifications, and justifications from students in an effort to make responses more conceptual in nature (Boaler & Staples, 2008; Hiebert & Wearne, 1993). The corresponding sub-construct of Student Providing considered the actual explanations of students followed from the press of Miss S or initiated by the students to clarify the mathematical contributions presented. For lesson one, the observation by the researcher was that Miss S had instances of missed press. The teacher, instead of pressing for a more conceptual response from students, would pose a question to another student, therefore limiting the possibilities for student providing situations. An example is shown in the excerpt below when talking about money value. The teacher asks the class to determine which has the greater value, 20¢ or $20. Student one says $20 is greater. The teacher then follows-up with another question for another student to state if the $20 is much larger in value or not. To receive the score for press, the teacher would need to follow-up with the same student.

Teacher: The counting is the same although the values are different. Which is bigger 20¢ or $20

Student 1: $20 (student name not captured because teacher pointed to the student)

Teacher: Is $20 a lot or a little bigger? (directed to student 2)

Student 2: a lot bigger

Teacher then moved on to another idea after the student responded. Here Miss S had a missed press recorded because she could have followed up with the student asking how he knew that $20 would be a lot bigger. There were three instances of missed press recorded by the researcher. Students were not required to back up their claims and the responses from Miss S's questions were in the form of one-word responses. Teacher press and Student Providing, therefore, received scores of two and one respectively.

Participation

Participation by the students was good. The teacher tried to include all students in answering questions throughout the lesson. Participation in the lesson refers to the percentage of students who made contributions as seen in figure 9. Therefore although the scores for other sub-constructs were rated low in some instances, students were required by Miss S to make contributions, and therefore, based on the rubric seen in figure 9, participation received a score of 4 since more than 75% of the students participated throughout the discussion.

Participation Rubric	
4	Over 75% of the students participated throughout the discussion.
3	50-75% of the students participated in the discussion.
2	25-50% of the students participated in the discussion.
1	Less than 25% of the students participated in the discussion.
0	None of the students participated in the discussion.
N/A	Reason:

Boston (2012). Used with permission of the author

Figure 9: Participation rubric ratings

The elements of each sub-construct and rating were perused and the rubrics that did not receive a score of four were compared to the researcher's notes. This was done to determine which areas specifically needed to be improved. For example the Potential of the Task received a score of three. The researcher determined that the task needed to include a requirement for students to explain their thinking. This was documented as an area to be discussed in the feedback session. The areas to be improved were then documented so that the feedback session would include them. From the IQA toolkit ratings for lesson one, the areas that were determined were the Potential of the Task, the implementation of the task and teacher and Student Linking. For the Potential of the Task students needed to be required to explain their thinking as a requirement to provide a complete response. For the implementation of the task, Miss S needed to reduce the level of assistance given to the students and students needed to present their ideas and explain to the class. There was also the need for Miss S to create opportunities where students were able to link ideas.

The Teacher Reflection

After each lesson, Miss S completed a reflection, with the aid of the reflection protocol provided (see Appendix D). In the reflection, she was prompted to reflect on areas that are included in the IQA toolkit such as the task given to students, class discussion, and student questioning. Miss S commented on her goal for the lesson, for students to count money of different money combinations. She stated that she noticed about three students who needed additional help after asking the class to give a 'thumbs up/thumbs down' to show if they understood. Students who identified with a 'thumbs down' were placed briefly in a small group and offered assistance by Miss S to address their misconceptions. Miss S stated their misconception concerned where to place the decimal point when writing money values. She also spoke about the class discussion and felt it was one of the best aspects of the lesson because she was able to link the concept she taught to the everyday lives of the students. Miss S also reflected on asking students which money value was greater and stated that students were able to complete the task with ease. Miss S felt that she asked the right number of questions of students and felt that overall the lesson went well.

After Miss S completed the reflection, the researcher read through the reflection to make notes of the parts of the lesson that Miss S considered important to mention in the reflection. This was done to be able to determine if there were any themes that arose from the reflection that could be related to the IQA toolkit.

For example there were two statements related to questioning from reflection 1:

Excerpt 1: "I asked several students if they would rather have twenty cents or twenty dollars in their pocket. They enjoyed this questioning and it helped them to realize the difference in value between coins and dollars"

Excerpt 2: "I felt that I asked the right amount of questions and tried to engage almost every student."

The reason excerpts related to questioning were noted is because questioning is a sub-construct of Academic Rigor. Therefore, since Miss S did not get a rating of four, the researcher noted contributions made by Miss S related to this sub-construct. Themes were determined from the statements made by Miss S in her reflection. Therefore when she made comments on the number of questions (excerpt two) asked of students and not the cognitive demand of the questions, it seemed important to be considered as a discussion point for the feedback session related to the theme of questioning. This inclusion in the feedback session was done in an effort to increase the focus on asking academically relevant questions in whole class discussions.

Feedback Session 1

Feedback session one occurred 11 days after the reflection was received from Miss S. The feedback sessions were held at the convenience of Miss S, who had various demands on her time as a result of internship and other personal issues. Feedback session one took place in the morning school session when students were at specials (activities with other teachers). Specials lasted for 40 minutes so the feedback session needed to be less than that time while allowing Miss S to also have some personal time. The first feedback session lasted approximately 20 minutes. In order to have reduced noise, the first feedback session was held in an unassigned

secluded room of the library. This allowed for the audio to be recorded with limited interruptions.

Following the analyses of the IQA toolkit results and Miss S' reflection, the researcher identified themes that emerged to determine how to approach the discussion for the feedback session. After the researcher acknowledged the themes that arose, the researcher then used the IQA toolkit rubric for the feedback sessions. The researcher focused on the qualitative explanations as a guide for discussing the themes. In preparation for the feedback session the researcher also decided on examples, using the observation notes as a guide, to use with Miss S to aid in her understanding of what was required for the lesson. For example, the concept of Teacher Press was an area that was discussed in feedback session one. Miss S did not ask students for extended explanations from students' contributions in the lesson. In the feedback session this was discussed with Miss S and the researcher used an example from the lesson, as shown in the excerpt below, to explain the concept of press.

Researcher: For students' response, press until you get a conceptual response from them.

Miss S: ok?

Researcher: For example when you asked students to tell whether $20 or 20¢ was bigger, a student said $20. You could have gone on to ask for example "how do you know $20 is bigger. If you ask 100¢ or $1 students should be able to respond. Sometimes you can use something else that is not on the paper.

Miss S: ok

Researcher: You use follow-up questions to build. This is important because students need to understand that the dollar amount is not necessarily less than the cents such as 100¢ and $1

Miss S: ok

Researcher: You can ask "what if I have this…? to see if they are able to respond. You can go on to build something from their answer.

Miss S: ok

Although the response from Miss S for this excerpt was "ok", she seemed to understand the idea of eliciting extended responses from students. She responded for other discussion points differently such as "I get that" or "that makes sense". Based on her response, audible and non-verbal gestures, the researcher was satisfied that she understood what was discussed.

To determine the aspects of the feedback session that would be pertinent to the study, the researcher watched the video as a whole. The entire video for each cycle was not transcribed because there were parts of the feedback discussion where Miss S mentioned information that was important to the general review of instruction but not necessarily important to the themes that would have been useful as they pertained to the lesson and constructs covered on the IQA toolkit. For example Miss S said two different statements: *"I think I did better than before with my questioning, but I know I need to continue to improve"* and *"my supervisor does not require me to write detailed plans but I do it anyway because it helps me".* Both these statements were made during the feedback session. However, only the first statement is directly related to instruction because questioning is an essential aspect of the lesson in engaging students. The researcher was therefore able to transcribe statements and phrases that had connections to the IQA toolkit and would therefore be imperative to the analysis of the feedback session.

The feedback session focused on sub-constructs of Academic Rigor and Accountable Talk. The researcher took Miss S through the lesson, asking questions where clarification was needed and discussing the important ideas related to the sub-constructs of Academic Rigor and Accountable Talk and discussed how to improve future lessons. The important references from feedback session one are noted in the following paragraphs.

The discussion begun with a look at the Potential of the Task. The nature of the task was discussed and the researcher noted that the task had good components but students were not required to explain their thinking. Examples of how to make the task more robust were discussed and Miss S' gestures suggested she understood what was said.

In discussing the implementation of the task, the focus was on Miss S and how she provided many answers for the students. It was also noted that she did not allow students to demonstrate enough to the rest of the class, thereby placing her in the position to demonstrate the processes, although observed from students' work. The researcher then went on to talk to Miss S, using examples from the lesson when she could have allowed students to be the ones leading the discussion.

Researcher: I don't know if you are aware of it and I do think many of us as teachers we are not aware of it, we speak a lot in our classes, and we may ask questions of students and still speak a lot. One thing I noticed is that you would ask students a question and say "do it with me"

Miss S: oh ok

Researcher: So you may want to ask those questions and give students the opportunity to respond, then you can reinforce after that

Miss S: ok, that makes sense

Researcher: So you may ask for example Josh count for me because when you ask students to answer together you may miss some students who are having difficulties.

Miss S: right ok, that makes sense

Researcher: When someone responds you may even ask another student "what did he say?" to see if they are listening and making sense of their classmates' understanding

Miss S: ok

Researcher: so pull back a little in terms of how much you help them. Give them an opportunity to see if they can work first. You say count and you…

Miss S: count with them (laughs). Ok

The previous excerpt follows from the section of the discussion where the researcher discussed with Miss S about allowing students to talk in the classroom. Miss S' response of 'that makes sense" may be determined as acceptance of the suggestion. She also agreed to her tendency to answer with the students instead of allowing them to answer on their own.

The participation from the students was good because Miss S allowed students, in some form, to make contributions to the lesson. The teacher was encouraged to continue including students but also to give opportunities for extended contributions. Questioning was another aspect that was discussed in the feedback session. The researcher discussed with Miss S her

questioning techniques and also the fact that she wrote in her reflection that she thought she asked the right number of questions.

Researcher: You said also that you thought you asked the right amount of question

Miss S: ok

Researcher: You did ask a number of questions but there are different types of questions. So there are questions that would result in a one-word response and there are questions that would result in an explanation. You want to balance questions that would require a one-word response and those that would require an explanation.

Miss S: ok

Researcher: Of course the questions that require an extended explanation are at a higher level.

Miss S: right, the higher order thinking

The researcher then gave Miss S an example of a higher-order question that she asked during the lesson to use as an example when compared to a question requiring a one-word response. It appeared that Miss S understood the differences in questioning type, not only from her stating the term higher-order questions, but from her response throughout the discussion.

Cycles 2 and 3

The researcher sought to identify the influence that feedback might have had on the instructional practices of Miss S. Therefore after instruction in the second cycle, comparisons of the themes discussed in the previous feedback session (session one) were analyzed to determine

if the feedback could have influenced any changes in instruction. This was again done in lesson three where comparisons were made and measured to lesson two using the Feedback Intervention Theory model to determine the possible influence of feedback on instructional practices. Following are lessons two and three, Miss S' reflections of the lessons, and descriptions of the feedback sessions

Lesson 2

In lesson two, students were required to measure the length of objects using a ruler. Miss S started out with a whole class discussion about the ruler. She presented two different rulers, A and B (Figure 10), to the students. Both rulers had inches on one side and centimeters on the other side with ruler B being wider in size. The focus of the teacher for this lesson was on measuring lengths in inches and therefore the inch measurement was emphasized.

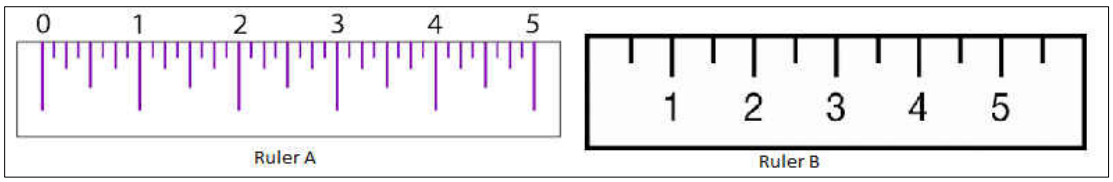


Image by author

Figure 10: Examples of rulers with different starting points

Ruler A had an indent before the measurement numbers were placed while Ruler B started measurements at the edge of the ruler. Miss S led students into a discussion about how both rulers were similar and how they were different. Students were able to state the similarities including that they both had lines, were used to measure, and the numbers went in order from smallest to largest. Students also observed differences such as the size of the rulers, the starting

point to measure for each would be different (one had an indent before starting at zero and the other did not), and different measurements (inches versus centimeters). Students were able to determine that the width of the ruler did not impact the measurement, stating that the rulers had the same measurement and number. Before the discussion, the teacher asked students to talk to their shoulder partners about the similarities and differences. The discussion also had students looking at the ruler that measures 12 inches and talking about if it was possible to measure objects longer than 12 inches.

Students were provided with a ruler to measure the lengths of some objects. The instructions for the task were projected along with examples of the objects students were asked to measure. The objects given such as a pencil and Popsicle sticks were not whole number measurements so Miss S told the students they had to estimate to the nearest whole number measurement. Students were required to share the materials but would measure on their own, documenting their answers in their books. The teacher engaged students in methods of estimation and students determined that the number the measurement is closest to would be used as the estimate of the length. An excerpt illustrating this discussion is provided below.

Teacher: Look at the Popsicle stick. Correctly measure the length of the stick. Write in your notebook what the length of the Popsicle stick is. Estimate to the closest inch.

Teacher: Show with your fingers the length you think the Popsicle stick is.

(Students stick their fingers in the air)

Teacher: I see 5 and 6. Rave, which is it?

Rave: 6 1/2. When you're counting and it's in the middle it's a half

Teacher: So is it between 5 and 6 or 6 and 7? Jody?

Jody: 5

Teacher: Measure to the closest inch

Jody: I accidentally said 5

Teacher: Grace?

Grace: 5 ½

Teacher: measure to the closest inch (without fractions)

Grace: 6

Teacher: Raise your hand if closest to 5. (Students raise hands)

Teacher: Raise your hand if closest to 6. (Students raise hand- majority)

Grace: For estimation you say closest to 6. There is a little gap to get to 6 and a bigger gap to get to 5.

Teacher: What if the Popsicle went here (between 8 and 9- closer to 9), how many inches would that be? Talk to your partner. What did you decide? Amanda

Amanda; 9 inches

Teacher: Did anyone else get that?

Class: yes

From the excerpt students were able to identify the estimated lengths of the different objects. They were also able to explain how they arrived at their estimated value. Noticeable is that the teacher called on students with different answers.

Students were given pieces of cord for them to find a partner with the same length of cord. Students walked around the class asking other students if they had cords of a specific length, such as 10 inches, so that the students could locate their partners. Those partners worked in pairs to consult each other on measuring the lengths of given objects. Students were then required to measure the actual lengths of other objects drawn in their text and then participate in the class discussion. The excerpt that follows is of a discussion involving Amy, and depicts a part of the discussion on estimating the measurement of the items students measured.

Teacher: How many inches was the paper clip? Amy?

Amy: three inches

Teacher: Was it right there on the 3 or between?

Amy: between the 2 and the 3

Teacher: How do you know?

Amy: It's much closer to the 3, farther away from 2

Miss S had similar conversations with at least 2 more students about measuring the lengths of various objects where they were required to estimate the lengths. The lesson ended with an enrichment activity from the text where students were required to measure the lengths of lines provided to the nearest inch. A brief discussion was held after the enrichment section of the

class although the period ended soon after. Miss S had a discussion with one student on how to measure the line as seen in the excerpt below.

Miss S: How do I measure Brit?

Brit: Measure each dot to the other. (Figure 11)

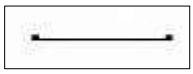


Image by author

Figure 11: Example of student's drawing

The student used the ruler to measure 1-inch spaces

Miss S: So you're saying I should write under. When I write what should I do?

Brit: put them all together

Miss S: What do I call adding over and over?

Brit: repeated addition

Miss S: Repeated addition is adding the same number. We are writing a number sentence.

The teacher reviewed the instructions with the students. However the class time had expired so the lesson was finished at this point for the day.

The IQA Toolkit Rubric Results

The method of determining the results from the IQA toolkit for lesson observation two was similar to that of lesson observation one. The observation notes were used as evidence to determine the descriptions from the rubric that corresponded with those descriptions and the scores were allotted based on those comparisons. Following are the results from the Academic Rigor and Accountable Talk sub-constructs for lesson two and a comparison to the sub-construct results for lesson observation one.

The Academic Rigor Rubrics

Table 6: Lesson 2 Academic Rigor rubric results

Sub-construct	IQA rubric rating
Potential of the Task	3
Implementation of the Task	4
Student Discussion after the Task	3
Questioning	4

Potential of the task

The task, although appropriate for the level of the students, did not explicitly ask students to show their thinking. Students were required to show their work and give explanations during the discussion segment of the lesson, but were not required to explain their thinking as a setup of the task. The task was not based on a procedure but students were asked to find how they would measure the lengths of objects. The Potential of the Task received similar ratings in both lessons, that of a three. In lesson two, similarly to lesson one, the task did not explicitly ask for students to explain their thinking

Implementation of the task

For lesson two the students used non-algorithmic thinking to measure the lengths of the objects given. The approach was not rehearsed and students provided their own ideas about how they could measure the lengths of objects (measuring directly with a ruler, drawing a line and measuring the line or drawing 1-inch marks and then counting how many marks there were). Students were allowed to decide on their approach to solving the problem using their choice of method and were able to make mathematical conclusions. In lesson two, the teacher allowed students to be more involved in the discussion and reduced how much she talked throughout the lesson. The rating for implementation for lesson two was a four compared to lesson one where the implementation received a score of three. In lesson one, during the implementation of the task, students were not provided with sufficient opportunities to display their ideas. In lesson two, however, the students were consistently provided with opportunities to discuss and explain their strategies and thinking.

The teacher asked students on a number of occasions to tell how they decided on their measurements. The excerpt below demonstrates a discussion sequence about the lengths students determined after they measured. Students were required to give estimations and Miss S asked for them to explain what they did.

The students were required to measure the length of a candy

Miss S: Miles measured a candy. Here is a picture of the candy.

Miss S: Kandy did Miles measure correctly?

Kandy: The edge of the candy should be at the edge of the ruler.

Miss S: Good. So if you measure it will be?

Kandy: Six inches.

Kandy: If you don't go to the zero, you will get it wrong (referring to her ruler).

There were at least eight recorded opportunities for students to share with the class. The talk from Miss S in lesson two was predominantly asking students to clarify their responses compared to lesson one where the teacher answered questions for the students.

Student Discussion following the Task

From the previous excerpts related to lesson two, it can be determined that students were given the opportunity to talk during whole class discussion. For lesson two students gave descriptions of the important ideas in solving the given task. Students provided explanations of why their strategies worked but often required press from Miss S to give better, although incomplete, responses. The Student Discussion for lesson two received a score of three while in lesson one this aspect of the lesson received a score of one. Differences observed between both lessons included students' discussions of their strategies and explaining the reasons they worked compared to lesson one where justifying of claims was not facilitated.

Questioning

The sub-construct of questioning for lesson two received a score of four. A score of fou4 means that there were at least three academically-relevant questions asked throughout the whole class discussion. For lesson two, there were occasions where the teacher asked appropriate

questions of the students to elicit responses that determined how much they understood an idea. Examples of instances where academically relevant questions were asked is demonstrated below.

(i) If there is no zero on the ruler, how do you determine where to begin your measurement?

This question was considered important because students needed to learn how to measure the length of an object if the number sequence does not begin at zero.

(ii) If the numbers were further apart on the ruler, would there be a difference in measurement?

This is a relevant question because students should be able to determine different measurements are found on the ruler, therefore some numbers would be further apart based on the unit used to measure. For example 2 cm and 2 inches would have the number 2 in different places on the ruler.

(iii) Does the size of the ruler make a difference in measurement (followed from comparing the width of rulers)

This question is important because students should recognize that the width of the ruler does not affect the length measurements so the measurement of length of a 12-inch ruler measures the same regardless of the width.

The Accountable Talk Rubrics

Table 7: Lesson 2 Academic Rigor rubric results

Sub-construct	IQA rubric rating
Teacher Press	2
Student Providing	2
Teacher Linking	2

Student Linking	2

Teacher Linking

For lesson two, there was not much effort made to create opportunities for students to connect their ideas. Therefore Teacher Linking received a score of two. In one instance the teacher asked the class "do you agree or disagree?" when a student responded to the question of where to start measuring on the ruler if the numbering does not begin at zero.

In lesson one there were at least five instances of revoicing, therefore lesson one received a score of one while in lesson two, there was greater evidence noted on the part of Miss S to create a link but only with the superficial question of "do you agree or disagree".

Student Linking

Student linking for lesson two received a score of two. Although there was no revoicing evidenced from students, there was one contribution where a student said, "I agree with Amy" when she answered a question by Miss S regarding the numbers that were not on a specific ruler shown. There was no effort recorded of the students trying to create any link with the contributions of Miss S or peers in lesson one. Lesson two, however, had one recorded instance of a link, therefore lesson two had a higher rating.

Teacher Press and Student Providing

Teacher Press in lesson two received a rating of two. In lesson two, there were two occurrences where the teacher attempted to ask students to extend their answers by providing justifications by asking, "How do you know?". Compared to lesson one where there were

instances of missed press, lesson two showed an improvement in this area. For Student Providing, students made efforts to provide evidence for their claims by responding appropriately to Miss S's question of "how do you know?". Students, when requested by the teacher, provided justifications for their claims, although sometimes incomplete justifications were given that required additional press.

Student Participation

Miss S for lesson two, similar to lesson one, required students to contribute to the lesson. The involvement of at least 75% of the class was again observed in lesson two. Therefore the score for participation remained at four for this sub-construct.

IQA Toolkit Themes for the Feedback Discussion

Similarly to post lesson one analysis, the IQA toolkit rubric was used as a reference for decisions regarding aspects of the researcher notes to be used as contributions to the feedback session. Similar themes arose as in lesson one from the IQA toolkit rubric such as the Potential of the Task, Implementation of the Task, Teacher and Student Linking. Teacher Press and Student Providing were also noted as topics to be discussed. As was described from lesson one, the sub-constructs that did not receive a score of four were integral in the discussion of the feedback. The teacher reflection was submitted after the lesson was taught. A description of the contents of the reflection follows.

Teacher Reflection 2

Miss S started out her reflection by stating what she taught for her lesson. She stated that the time was reduced by approximately 10 minutes. On the day she was observed, school ended

early, thereby shortening instruction time. She mentioned, however, that she tried to implement the feedback she received from the previous lesson observation. Miss S said she felt the lesson was an improvement, although the strategies from the feedback she received would be developed over time.

Miss S then went on to speak about her lesson and what she required her students to do, giving a synopsis of the lesson. Miss S went on to talk about her questions stating that she thought the questions asked of students were 'academically relevant', actually mentioning the term.

"I believe that the questions I asked students were academically relevant. I made sure to ask questions that would help students to better understand the learning goal and standard for the lesson"

Miss S explained that she gave each person the chance to be actively involved in the lesson because she also allowed students to talk with their classmates. She went on to indicate that she felt the discussion was effective and that she asked the appropriate number of questions.

"I felt that the initial discussion was effective for students and that appropriate amount of questioning was provided. In comparison to my previous observed lesson I felt that I did a better job allowing the students to talk in order to facilitate their own learning."

Miss S also made a comparison between lessons one and two and felt she did a better job allowing students to talk in order to facilitate their own learning. Miss S noted that she has room to improve related to higher order thinking questions and wants to make it a natural part of her teaching.

"I feel that I still have room for improvement when it comes to giving students higher order thinking questions. I want this to be something that comes natural to me as an educator because I genuinely feel it has a high impact on student learning".

The Miss S's reflection brought up themes such as lesson implementation and student questioning. These themes were similar to the themes that were determined from the IQA toolkit that were to be discussed in the reflection and no additional themes were determined from the reflection.

Feedback Session 2

Feedback session two occurred three days after the reflection was received from Miss S. Feedback session two, held in the morning session, lasted approximately 20 minutes. The session was held in Miss S's classroom that was available because students were again at specials. During this feedback session, Miss S took notes from the session, a noticeably different approach from feedback session one. The themes that arose from the IQA toolkit analysis and the teacher reflection were discussed in the feedback session.

The session begun with the researcher talking about the task that the students were given. The researcher drew Miss S's attention to the fact that the task did not explicitly ask students to document their thinking and discussed with Miss S how to make the task a higher quality. The researcher also explained that the task setup does not only include what is written but accompanying requirements from the teacher are also considered.

Researcher: I remember the last time we spoke about the task.

Miss S: Yes

Researcher: Remember that if the task does not explicitly state that students should explain their thinking, you can have a good task but…

Miss S: not at the level you need it to be.

Researcher: yes

Researcher: The task as it is setup is not necessarily what's on paper, which is why I write everything you say, but also what you require students to do with the task.

Miss S: ok

The questioning aspect of the lesson was also discussed and the researcher expressed to Miss S that there was an improvement in the level of questioning to asking more higher-order-thinking questions that required students to think in order to respond with more than single-word responses.

Researcher: You had really good questions.

Miss S: Thank you.

Researcher: Not just "what" questions but you were asking students about numbers. Although it was a simple item as a ruler you were asking if the numbers were evenly spaced, or not evenly spaced would it make a difference or if there is not a zero on the ruler how to measure, because those are important questions to ask and not just assume students will know. I listened to your line of questioning and students' responses and thought it was pretty good

Miss S: Cool.

The researcher also spoke about the activities that students were given so they could have produced work to present. During the presentation of the students, the teacher's linking was also observed and the instance of a link was used as an example in the feedback session. The researcher then went on to give an example of a link using the students' work. Miss S was seen taking notes specifically for this segment related to student and Teacher Linking.

Researcher: In terms of Teacher Linking, remember that linking refers to how is it that you get students to link their ideas or to yours. I saw one example of a link.

Miss S: ok

Researcher: It was weak but when you said Aly's work is also two inches from the box, when compared to Ray's. So you made some connection with Aly's work and Ray's.

Miss S: ok

Researcher: You made some connection there by saying this person's work is similar but she did it vertically and he did it horizontally.

Miss S: ok

Researcher: You made the comment they are similar. You want to do more of that, so if you have more than one student presenting you want to make some link with their ideas, especially if they are different ideas. You also want to choose students with different ideas to come to the board, which are not incorrect responses necessarily but they have chosen a different approach

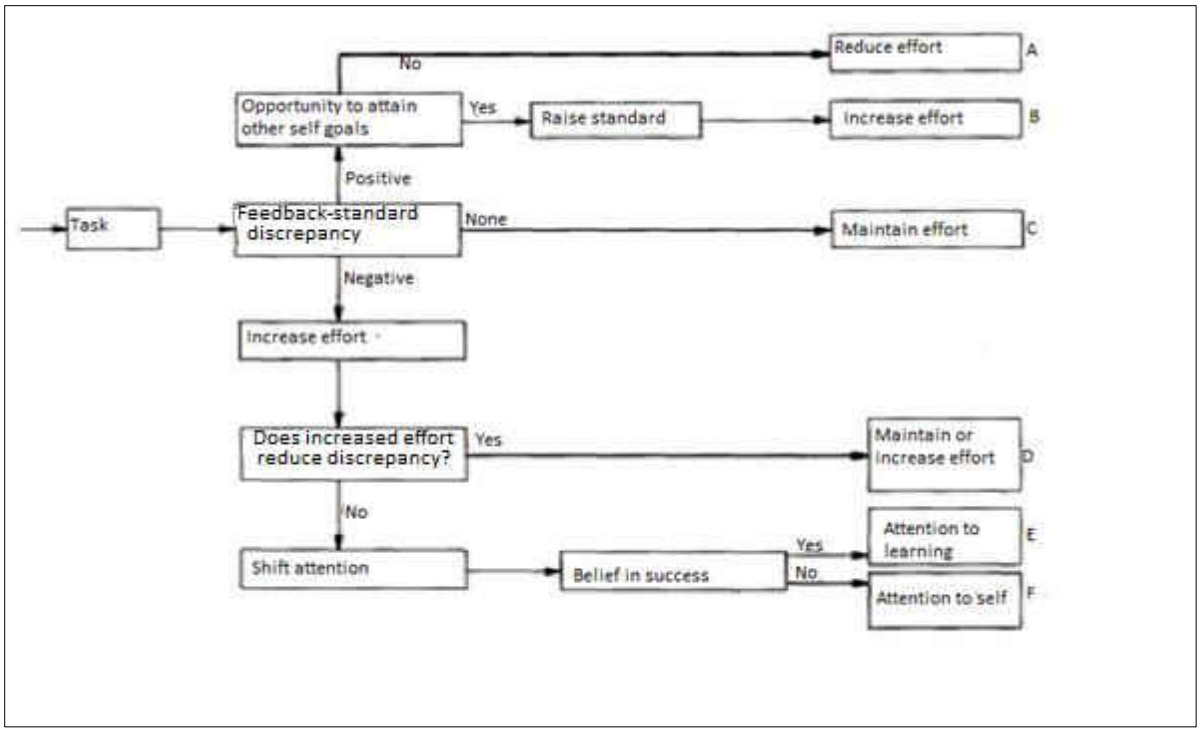
Miss S: ok

Researcher: and that lends your class to look at for example Ray did it this way and Aly did it this way. You can ask how are they alike, how are they different? And that's the link you can create for your students.

This discussion sequence highlighted the ways in which Miss S could have improved her lesson. It was noted that she took personal notes also from this discussion sequence, possibly providing examples for her next lesson. This area of the discussion was of interest because the researcher was curious of the nature of any modifications that would be done by Miss S.

Comparison of lessons one and two and the Feedback Intervention Theory (FIT) Connection

The feedback Intervention Theory (Kluger & DeNisi, 1996) represents feedback intervention and the possible responses to feedback. For this study, the nature of the response to feedback given was captured using figure 12. There are 6 possible response outcomes, labeled A to F. If there was a positive feedback provided to Miss S, the follow-up responses were captured in response A and B and if the feedback was negative, the response was captured from C to F.



Reprinted from "The effects of feedback interventions on performance: A historical review, a meta-analysis and a preliminary feedback intervention theory," by A.N. Kluger and A. DeNisi, 1996, *Psychological Bulletin*, p.264. Copyright 1996 by American Psychological Association publications

Figure 12: The effects of feedback intervention (FI) induced attention on task motivation processes and their consequences (lettered)

To determine the response of Miss S, observations one and two were compared. The comparison focused on each item discussed in the feedback, whether it was positive or negative. Once this comparison was determined, the IQA toolkit score was examined to determine the type of response exhibited by Miss S in the subsequent lesson.

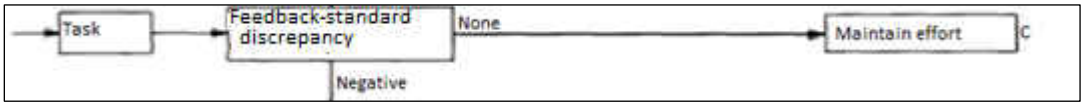
Responses to feedback Items

In the first feedback session the items related to the IQA toolkit that were discussed included the Potential of the Task, implementation the task, pressing for conceptual responses from students, and linking on the parts of the students and teacher. Table 8 summarizes the progressions between lessons one and two.

Table 8: Comparison of IQA results and FIT output

Feedback Discussion Item	Score for Lesson 1	Score for Lesson 2	FIT output
Potential of the Task	3	3	C
Implementation of the Task	3	4	D
Questioning	1	4	D
Student Discussion after the Task	2	3	D
Teacher Press	1	2	D
Student Providing	1	2	D
Teacher Linking	1	2	D
Student Linking	1	2	D

The Potential of the Task was given a score of three. In the feedback session, Miss S was made aware that the score was not the maximum possible, which would be considered negative feedback. Although Miss S was commended for her effort and received a score of three, it is considered negative feedback because there was improvement that was needed on the part of Miss S to make the task more robust for students by requiring them to explain their thinking. The method of making the task more robust was discussed with Miss S along with the steps that she could take to achieve this. However, in the second lesson, a similar level task was given to the students, which resulted in a score of three again. This meant that according to the FIT, Miss S received a response outcome of C because no change in action was taken and therefore the result remained the same (Figure 13).

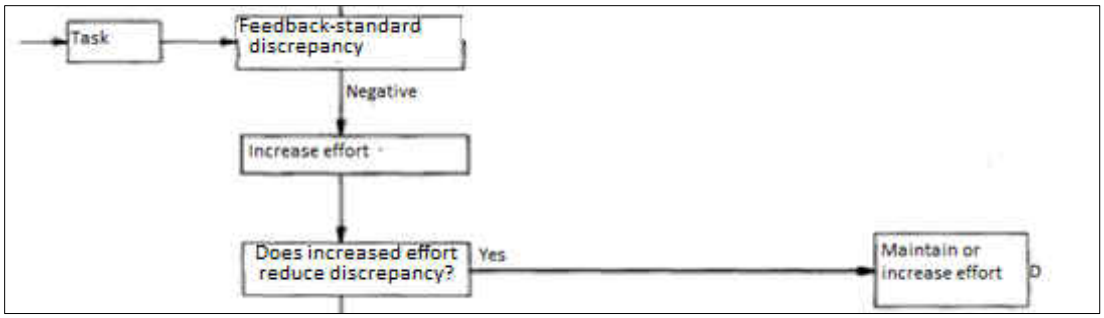


Adapted from "The effects of feedback interventions on performance: A historical review, a meta-analysis and a preliminary feedback intervention theory," by A.N. Kluger and A. DeNisi, 1996, *Psychological Bulletin*, p.264. Copyright 1996 by American Psychological Association publications

Figure 13: The effects of feedback intervention- Output C

In the first feedback session, the implementation of the task was also considered. In the first observation, Miss S received a score of three. The discussions related to implementation included the engagement of students. Although students were engaged, Miss S talked extensively during the lesson. She also needed to focus more on the depth of understanding of the mathematical concept and also to ensure that students were making mathematical

connections, with others work and with the mathematics itself. The feedback from lesson one was considered negative and therefore the response of Miss S had the options of remaining at the same level, improving, or showing a decrease in the level of performance. The second lesson score for implementation, however, was a four. There was much improvement noted from lesson one to lesson two as it related to the implementation. The students were more involved in the lesson, using their ideas to connect to others and demonstrating understanding of the mathematics. Since lesson two demonstrated an improvement, on the FIT model, corresponding with D, which indicated that the feedback was negative and that increased effort was made that reduced the discrepancy as demonstrated in figure 14.



Adapted from "The effects of feedback interventions on performance: A historical review, a meta-analysis and a preliminary feedback intervention theory," by A.N. Kluger and A. DeNisi, 1996, *Psychological Bulletin*, p.264. Copyright 1996 by American Psychological Association publications

Figure 14: The effects of feedback intervention- Output D

The discussion score increased from a two to a three. In the feedback session, discussion surrounding the rubric score focused on students being active participants in the discussion process, sharing their work, explaining their thinking and taking steps to understand the thinking

of others. There was an increased effort noted after the discussion session, on the part of Miss S to engage students in explaining their thinking and responding to higher levels of mathematical explanations and therefore the feedback would have been considered negative. In observing the lesson and scoring the IQA toolkit rubric, there was an improved effort from Miss S to have a meaningful whole class discussion where students presented their work and explained their methods to the class as required. The involvement of students in the discussion in a meaningful way resulted in an increased score for the discussion as measured by the IQA toolkit rubric. Corresponding to the FIT, the response of Miss S is recorded similarly to the implementation where there was an increased effort that resulted in a reduced discrepancy represented by D shown in figure 14.

The Accountable Talk sub-constructs were also compared to the FIT model. For Teacher Press, there was an increase from a rating of one for lesson one to a rating of two for lesson two. In Lesson two there was one instance of press by Miss S recorded which was not observed in lesson one. Feedback session one provided examples of press to Miss S for her to be able to require more conceptual responses from students by asking follow-up questions. An increase in score would suggest that there was possible effort made by Miss S to improve the level of responses from students in the second lesson. Relating the changes from lesson one to lesson two according to the FIT model, Miss S would have received negative feedback but showed some improvement in the effort resulting in increased performance, as shown by the D in figure 14. Student Providing rubric also received an increased rating, potentially because the press from Miss S prompted students to provide more conceptual responses, which they did.

The teacher and students' linking rubrics recorded improvements in the IQA toolkit rating. The teacher asked the question of "do you agree or disagree?" and in one instance a student said, "I agree with Max". Although the effort to link was minimal on both the part of the teacher and students, the increase in effort was noteworthy to increase the rating. In feedback session one, discussions were held with Miss S about how to help students to create links with the mathematics contributions of teacher and peers. The improvement on the negative feedback would also be represented by the D model (figure 14) of the FIT diagram.

Observation 3

The lesson in observation three took place three weeks after observation two and lasted for approximately one hour. The class consisted of 14 students since the other students were taken out by the intervention specialist. The supervising teacher was present for the duration of the lesson and had more input than in the previously-observed lessons, often saying "remember to explain your thinking" while students were on task. Students remained at their assigned seats for the duration of the class and talked to their shoulder partners when required by the teacher.

The concept that was taught was a continuation of the measurement unit that focused on metric measurements. Miss S started the lesson by asking students the units used to measure length. Students responded by stating inches and colored tiles. Miss S went ahead to point out that the rulers they have been using also has another unit of measurement, that of centimeters. Students were provided with 1-cm cubes that were used to measure pieces of yarn provided. Students were asked to explain the strategy they used to determine the total length of the pieces of yarn. The strategies included putting cubes side by side and counting the number of cubes.

Students also drew a line, then placed cubes along the line and counted the number of cubes to measure the length in centimeters.

The teacher created the opportunity for students to make a link with another student's work consistently throughout the discussion. She asked students "who did it another way?" or "who did it a similar way?" Students would respond or say without prompt " I did it like..…" referring to their classmates' names or "I checked mine with _____ strategy". One student said she used a combination of two strategies by two other students to create one of her strategies.

In lesson three, students were actively involved in the lesson, contributing their methods and their ideas and making links with the work of their peers. Students shared their work, by explaining to the class and responding to Miss S's presses to provide more conceptual responses than were initially given. The results according to the IQA toolkit rubric and comparison to lesson two follow.

The IQA Toolkit Rubric Results

The Academic Rigor Rubrics

Table 9: Lesson 3 Academic Rigor rubric results

Sub-construct	IQA rubric rating
Potential of the Task	4
Implementation of the Task	4
Student Discussion after the Task	4
Questioning	3

Potential of the Task

For the Potential of the Task, students were given a task for which they needed to develop strategies and explain their thinking as it related to their solution processes. Miss S consistently reminded students that they needed to explain their thinking throughout the process of completing the task. The rating in lesson three compared to lesson four showed an increase in the score because in lesson three, the task was genuine and required students to develop strategies to solve using their ideas and also required students to explain their thinking. The task potential for lesson three receive a four, the highest rating of the three observed lessons.

Implementation of the Task

Task implementation for lesson three received a rating of four. In lesson three, the students were given numerous opportunities to be involved in the lesson by explaining their solutions and their thinking processes and methods for solving the task. Students were also required to justify their responses consistently with appropriate mathematical explanations. For lesson three, Miss S noticeably did less talking but followed up students' contributions with prompts to help them provide conceptual responses. Students made meaningful connections in both lessons. Lesson three, however had more deliberate instances for students to make meaningful mathematical connections. The improved implementation strategies from Miss S resulted in an increased rating for Implementation of the Lesson rubric.

Student Discussion following the Task

The class discussion in lesson three was robust with students contributing their ideas, methods, and work. Students were led into presenting their work and justifying their claims in

numerous instances throughout the lesson. The rich discussion in which the students were involved produced an IQA toolkit rating of four. Students presented extended responses for their claims and represented more than one strategy for their solutions. In lesson two, Student Discussion sub-construct received a rating of three because students often required extended press from the teacher to provide conceptual responses and the discussion was not as robust as in observation three.

Questioning

In lesson three, the teacher asked academically relevant questions of students. There were two instances where the teacher asked these questions. During the lesson, the teacher asked a student who measured the length of a marker with the 1-cm cubes to measure the length of a crayon. The student took all the cubes he used to measure the marker and lined it up against the crayon. He took away the cubes one at a time until the rest had lined up with the crayon. Miss S then asked, "If the marker was 12 cm and the crayon is 8 cm, how many cubes did Brit push away?" Although the student followed up with a response of four, the question required students to look at the relationships between the two quantities to respond. In the other instance the talk surrounded the way in which students chose to line up the cubes. Students discussed what would happen if the cubes did not fit to each other. Miss S asked if the cubes were not lined up perfectly, if the measurement would be higher or lower. A student responded by saying that if they were not lined up correctly, then that would result in a higher number because of the spaces between the cubes. The rating for lesson two for questioning was four which reduced in lesson three with a three because of the number of instances academically relevant questions was asked.

The Accountable Talk Rubrics

Table 10: Lesson 3 Accountable Talk rubric results

Sub-construct	IQA rubric rating
Teacher Press	3
Student Providing	3
Teacher Linking	3
Student Linking	3

Teacher Linking and Student Linking

Teacher linking in lesson three was more consistent than in lesson two. Lesson three had consistent links created by the teacher throughout the class discussion. The teacher consistently asked throughout the discussion if anyone solved the problems in similar or different ways. There were also 10 instances of teacher revoicing noted by the researcher during the lesson. Compared to lesson two with one instance of a link noted, lesson three, had the teacher creating more opportunities for students to link ideas.

The teacher when assigning work to students said, "*go ahead and measure number 3. See if you can use 2 different strategies: one of your own and one from someone else"*. Another time during the lesson Miss S asked, "*did anyone do the task a different way?"* and called on Grey to answer. He responded by saying, "*just to check, I did what Mary did, I used my number line*". Teacher Linking and Student Linking both received a score of three, while in lesson two they received scores of two each.

Teacher press and Student Providing

Although there was more effort from students to explain their solution methods, Miss S often required additional information from them to achieve more conceptual responses. There was more press required from students for lesson two because in lesson three, students provided more information on their own and required less press from Miss S when compared to lesson two. In lesson three, the rubric rating was three for Teacher Press and Student Providing compared to lesson two where the ratings was two for both sub-constructs.

Teacher Reflection 3

Miss S started her reflection by providing a synopsis of the lesson. She then described the stages of her lesson, that she required students to measure pieces of yarn, and that students responded by saying they could measure length in inches and with colored tiles. She explained that she told students about another measurement on the ruler, that of centimeters. Miss S went on to describe that students had several strategies to measure the lengths given and that she included in her discussion that several strategies can lead to the same result". Miss S spoke of allowing students to describe their work to their peers and that she tried to relate strategies to others they had been introduced to before.

Miss S spoke of her rich task and that she consistently asked students to explain their thinking and strategy. She finished off her reflection by stating that she had improved her questioning skills related to pushing students to share their thinking and prove their answers.

Feedback Session 3

Feedback session three was held seven days after the reflection was received from Miss S. The venue for the session was the same as session two, again because the classroom was available. Since the feedback session occurred closer to the end of the placement, it became increasingly difficult to schedule sessions. Miss S, however, had some available time that could be used for the session. The session lasted approximately 12 minutes. The nature of feedback session three was a little different from the previous sessions, because discussion was focused on lesson three, as well as looking back at previous lessons and talking about how elements of the lessons had evolved over the course of the observations.

The researcher reviewed the elements that Miss S reflected on as well as the IQA toolkit results. Discussions were held on the task, how it was implemented, questioning, linking, and press situations. The researcher and Miss S spoke about how students were given the opportunity to think, work on their own, share their work, and explain their thinking. Additionally, the researcher reviewed the progress the teacher had made over the period by using lesson 3 as an example, when compared to lesson one. As part of Miss S's internship requirements, she was expected to observe other teachers. The researcher encouraged Miss S to think about the elements that were discussed during the feedback session and look for those elements in the lesson of the classroom teacher. She was also encouraged to keep improving in her instructional practices.

Comparison of lesson three to previous lessons and the FIT

The changes in score from lessons two to three can also be represented on the FIT diagram. Table 11 represents the differences in scores from lessons two to three and how these scores correspond with the FIT output in table 11. There were instances of both an increase and

decrease in scores between the two lessons. Reference to the FIT output may be found in figure 12.

Table 11: Comparison of IQA results and FIT output

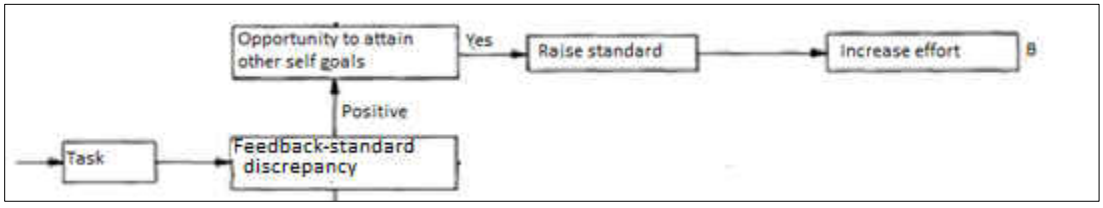
Feedback Discussion Item	Score for Lesson 2	Score for Lesson 3	FIT output
Potential of the Task	3	4	D
Implementation of the Task	4	4	B
Student Discussion after the Task	3	4	D
Questioning	4	3	A
Teacher Press	2	3	D
Teacher Linking	2	3	D
Student Linking	2	3	D

The Potential of the Task improved to a four from a three in lesson two. For this task students were asked to explain their thinking, which was not observed in lessons one and two. The Potential of the Task, then received a score of four, having met the requirement by the IQA toolkit to be rated at the maximum score possible. The feedback would have been negative for the Potential of the Task in feedback session two but was improved in lesson three. According to the FIT diagram, Miss S would have a feedback intervention output represented by D in figure 14.

Implementation of the Task

The implementation of the task from lesson two to lesson three maintained the same IQA rating. In the feedback, the response would have been considered to be positive and would have to compare lesson three. Lesson three, although receiving the same rating from the IQA toolkit as lesson two, provided more opportunities for students to demonstrate their thinking while understanding the mathematical thinking of others.

Based on the FIT, a positive feedback may have yielded more than one result. In reflection three, it could be seen where Miss S stated that she improved her skills, which may indicate that she gained from the feedback process. Connecting with what Miss S said, and based on improvements in lesson implementation, although the score did not change, the FIT process connecting to this idea would be B as demonstrated in figure 15 since there was an increase in effort as a result.



Adapted from "The effects of feedback interventions on performance: A historical review, a meta-analysis and a preliminary feedback intervention theory," by A.N. Kluger and A. DeNisi, 1996, *Psychological Bulletin*, p.264. Copyright 1996 by American Psychological Association publications

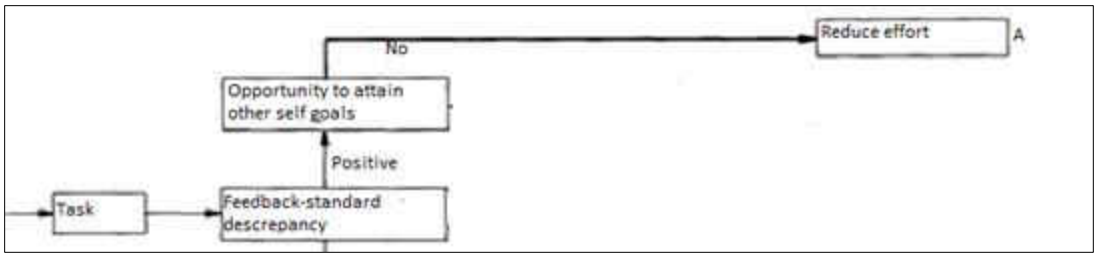
Figure 15: The effects of feedback intervention- Output B

Student Discussion

In observation three, the Student Discussion after the Task had an increased score to a four. In the third lesson, there was more meaningful involvement from the students as it related to sharing their methods and ideas with the class. The teacher better directed the students towards description of their work and justifying their contributions, with mathematically acceptable responses. Observation two received a score of three for Student Discussion, so there was an improvement in the discussion level of the lesson. The feedback intervention output for this sub-construct would be D represented by figure 14.

Questioning

Questioning for lesson three reduced in the level of rating from lesson two. Students were asked two academically relevant questions to demonstrate their understanding of the concept taught. In lesson two, academically relevant questions were asked of the students an appropriate number of times as measured by the IQA toolkit. Positive feedback was received, but Miss S did not ask enough academically relevant questions in her follow-up lesson, so the rating was reduced. According to the feedback intervention output where a positive feedback is received and there is a reduction of effort in the follow-up, the feedback output would be represented with the letter A (see figure 16).



Adapted from "The effects of feedback interventions on performance: A historical review, a meta-analysis and a preliminary feedback intervention theory," by A.N. Kluger and A. DeNisi, 1996, *Psychological Bulletin*, p.264. Copyright 1996 by American Psychological Association publications

Figure 16: The effects of feedback intervention- Output A

Teacher Linking and Student Linking

The linking rubric for Miss S and students both had increased ratings. There was noted effort on the part of Miss S to create opportunities for the students to link their ideas to each other. Similarly, the students demonstrated the ability to link ideas and methods with that of their peers. An increase in these two areas after feedback would correspond with the FIT process of D.

Teacher Press and Student Providing

The press and providing rubrics also showed improvements from lesson two to three. Miss S, in lesson three, required students to justify their responses by answering "why" questions to justify their claims. Students followed with providing responses that clarified their

contributions twice during the lesson. This process and results, reflecting improvement in ratings correspond to FIT level A as seen in **figure 16**.

Miss S' Observation of her Supervising Teacher

Miss S had observed her assigned teacher early in the semester as a requirement of her internship in preparation to teach, but did not have detailed notes of the observation. However, she was requested by the researcher to observe her supervising teacher again following the third feedback session. For the purpose of the research Miss S observed her assigned teacher teaching mathematics once and wrote a reflection on what she observed. The observation was requested by the researcher to determine if Miss S was able to identify or notice aspects of the lesson that would be captured by the IQA toolkit, specifically those areas that the feedback sessions focused on. Miss S observed the lesson of her supervising teacher after her three feedback cycles were completed.

The lesson Miss S observed focused on using expanded form to solve three-digit addition problems. Miss S commented on the Potential of the Task the students were given. She stated that, "*my supervising teacher provided a rich task for her students*". The only contribution in her reflection that gave support to her claim of a rich task was she mentioned that, "*students were given challenging tasks*". She also stated that students displayed multiple approaches to solve the tasks, which may suggest that the tasks did not present predetermined steps for reaching solutions. Miss S provided additional support for her observation by sharing examples of interactions between the teacher and students. When students were asked to work on the task the teacher said, "*okay now look at problem 1. Try to work this problem out on your own using expanded form or any other strategy you think might work*".

Students were asked to share their approaches to solve the problems to the class and students who were seated were asked to follow on their white boards to try to understand or make sense of their classmates' contributions. Some students used a combination of more than one strategy to solve tasks as well. For the lesson she observed, Miss S stated that she saw the teacher use *Student Linking*. For the contributions that would be considered as linking, Miss S stated that students, with the teacher's help, were able to find other students with similar reasoning as well as those who solved the problems differently. After one student presented a strategy, the teacher asked, "who used a different strategy?" and asked a student to share the strategy with the class. Miss S also noted that the supervising teacher provided multiple opportunities for students to *explain their thinking* and share their strategies with the class. Miss S also mentioned she observed that the teacher *did little talking and when she did, she connected students' ideas*.

The observations Miss S noted are consistent with the descriptions that are captured by some of the sub-constructs of the IQA toolkit such as the Potential of the Task, student and Teacher Linking, whole class discussion sessions, and students being involved in the class by explaining their ideas and strategies. Miss S said she benefitted from the overall process of observation and feedback, and would continue to implement the strategies learned through participating in the research.

The data analysis in chapter four served as an important aspect to answering the research questions focused on determining if feedback on mathematics instruction, using the IQA toolkit has any influence on the instructional practice of a preservice elementary teacher. Chapter five provides a response to the research questions and continues with a look into what the analysis

means and how this research may inform decisions made regarding clinical experiences in mathematics.

CHAPTER 5: DISCUSSION

Purpose and Overview of Methodology

The purpose of this study was to determine if feedback influenced the instructional quality in mathematics of a preservice elementary teacher. The study was conducted during the PST's final internship and utilized the Instructional Quality Assessment toolkit (Boston, 2012) as a guide to the purposeful feedback that was provided to the participant. The IQA toolkit contains qualitative descriptions that are related to classroom behaviors, on the part of the teacher and her students. The behaviors described in the IQA toolkit were compared to the observed behaviors during instruction to provide a score. The participant also reflected on her lesson, which provided data on her thoughts on the effectiveness of her instruction. These thoughts were compared to the categories on the IQA toolkit to determine if there were any connections to these categories.

To conduct analysis the sub-constructs from Academic Rigor and Accountable Talk of the IQA toolkit were used as themes and the lesson was analyzed to determine how it compared to the themes from the sub-constructs. The lesson reflections of the participant were also analyzed to determine the elements of the reflection that corresponded with elements of the sub-constructs from the IQA toolkit. This was compared for each lesson to determine the influence feedback may have had on the instructional quality of the participant.

Additionally, the participant was requested to observe her supervising teacher without further direction beyond that she should observe and record her reflections, after the three cycles of the iterative process of teaching, feedback, and teaching based on feedback, in which she was

involved was completed. The participant identified, based on her observations, connections to the IQA toolkit rubrics that were used in the feedback phase of the process. Essential to the discussion provided here is the ability of the participant to identify, or notice, the connections between the feedback she had received on her own teaching and the quality of her observation of the instruction of someone else, that of her supervising teacher.

Summary and Discussion of Key Findings

This chapter starts with a summary and discussion of the findings related to the research organized by research questions. The implications of the study to policy and practice in relation to the evaluation in mathematics for PSTs are discussed. It continues with discussion on the concept of noticing that emerged during the research. The participant made instructional moves during the lesson in response to students' contributions, which can be argued to be influenced by teacher noticing. Additionally, the participant was asked to observe her supervising teacher and write a reflection on her observation. Comparisons or connections of the lesson made by the participant to the sub-constructs of the IQA toolkit, stemmed from the participant making interpretations from her observation and reflecting on the instructional quality of the supervising teacher by making connections to the tool. The limitations of the study are also discussed and finally the potential contributions to future research are presented.

Discussion of findings

The main research question to which two sub-questions are related is: How does feedback using the Instructional Quality Assessment (IQA) toolkit influence the mathematics instructional practices of a final year preservice elementary teacher? Both sub-questions include

specifically the constructs of Academic Rigor and Accountable Talk and how feedback may have influenced the elements of instruction identified by these constructs. The participant's instruction reflected in the results from the IQA toolkit as well as her reflection on instruction provided evidence of the influence that feedback may have had on her instructional quality. The answers to the research questions painted a picture of the noticeable differences observed in the instruction of the participant, Miss S, over a period of three instruction sequences described by an iterative process of teaching, feedback, and teaching based on feedback.

Research Question 1

The first research sub-question was: In what ways does feedback using the construct of Academic Rigor as measured by the Instructional Quality Assessment toolkit influence the mathematics instructional quality of a final year preservice elementary teacher? An analysis of research question one yielded findings as it related to feedback's influence on instructional quality using the IQA toolkit.

The results of the IQA toolkit evaluation comparing lessons one and two and lessons two and three indicated differences in IQA toolkit rubric scores in both instances. An examination of the scores from the three observations showed that no score in lessons two and three was less than the score given for lesson one for the Potential of the task, Implementation of the Task, Questioning, or Student Discussion after the Task. After each feedback session, the scores for the next lesson remained the same or increased when compared to lesson one. For feedback session one, all the sub-constructs were discussed. In lesson two, all of the four sub-constructs discussed had an improvement in the rating of the instructional sequence. Therefore, after the first feedback session, all of the sub-constructs showed improvement in ratings. However, the questioning rating was reduced in lesson 3.

For the Potential of the Task, the influence was determined from the second feedback session because the score for this sub-construct remained the same for lesson two as for lesson one, that is, both lessons one and two received a score of three and lesson three received a score of four. The requirement for the tasks was the same for the first two lessons, so there was no change in score. For lesson three, the students were required to explain their thinking by providing explanations and justifications which resulted in a higher score from the IQA toolkit rubric results. During feedback sessions one and two, the researcher discussed how the Potential of the Task could be increased.

Student Discussion after the Task also recorded differences in score from the IQA toolkit rubric. The score after feedback session one increased for this sub-construct from a score of one to a score of three. The difference in the discussion was observed by students being required to present their strategies for solving the task as well as explaining the reasons their strategies worked. Press was also required for students to extend their responses to provide answers that were more conceptual. Discussion for lesson three revealed even better results following feedback session two. Discussion for lesson three received a score of four because of the rich talk with which the students were engaged by presenting their work and justifying their claims consistently throughout the discussion surrounding the task.

Academically-relevant questions were asked in at least three instances in lesson two when compared to lesson one. The presence of these questions increased the score after feedback session one. However, in lesson three the score decreased because, although there were academically-relevant questions, there were not enough for a score of four to be given. Miss S noted this area as one she personally needed to work on to ensure that these types of questions are asked to students during her instruction.

Corresponding to the IQA toolkit rubric are the areas that the participant discussed or mentioned in some form in the reflection. In reflection one she talked about her questioning techniques and opportunities provided to students during instruction. From reflection one, one point was directly related to the Academic Rigor rubrics, that of questioning. In reflection two, Miss S also spoke of questioning but also included the term "academically relevant", which is somewhat unique in the qualitative descriptions in the Academic Rigor sub-construct of Questioning. Additionally she spoke about allowing students to talk with each other and facilitating their own learning. In reflecting on lesson two, the participant went into more detail than in reflection one and therefore the researcher was better able to determine the connections between her reflection and the themes from the IQA toolkit. In reflection three, notably the participant mentioned her "rich task", that was not mentioned in that form in previous reflections. She also spoke about her own abilities to ask more academically-relevant questions.

It can be determined there was influence on instruction based upon the feedback. Apart from the participant mentioning that she tried to implement the strategies, the follow-up lesson saw improved scores in a number of instances from the Academic Rigor rubrics, demonstrating a number of changes during her instruction as she sought to engage students. The language in the reflections also showed differences. It can be determined that influences from feedback in Academic Rigor can be seen from the improvement in the instructional quality reflected in the increase in rubric scores over the lesson observations. Influence may also be determined from the language the participant used such as "academically relevant" when relating to questions, "rich task" when talking about the task, or her focus on what is considered important in a lesson shifting from superficial importance such as asking many questions to focusing on asking higher-order questions. Her implementation of the lesson also changed, allowing multiple solution

strategies among students, and facilitating students' interactions during instruction. The tasks that were given to students also showed change from asking students to simply solve the problem, to asking students to explain their thinking, document their solutions, and justify their methods.

The influences identified can be linked to areas of the IQA toolkit, specifically the construct of Academic Rigor. It can therefore be concluded that there was influence on the instructional quality of the preservice teacher as it relates to the construct of Academic Rigor. Additionally, the iterative process of teaching, feedback, and teaching based on feedback using the IQA toolkit served as an avenue to improve the quality of instruction by providing focused feedback.

Research Question 2

The second research sub-question was: In what ways does feedback using the construct of Accountable Talk as measured by the Instructional Quality Assessment toolkit influence the mathematics instructional quality of a final year preservice elementary teacher? Similarly to Academic Rigor, there were perceived influences that feedback had on the sub-constructs of Teacher Linking, Student Linking, Teacher Press, and Student Providing, which improved instructional quality. For feedback session one, there was discussion with the participant regarding Teacher and Student Linking, as well as Teacher Press. From observation two, there was an increase in score for each of the sub-constructs of Accountable Talk when compared to lesson one, though not achieving the maximum score possible.

It could therefore be determined that there was an influence on the instructional quality related to the Accountable Talk Rubrics from the first feedback. The influence on instruction as a result of the feedback would be determined by the differences observed between lessons one

and two and lessons two and three. Discussion in feedback session two was connected to each of the sub-rubrics for Accountable Talk. The results for lesson three from the IQA toolkit rating showed increases in Teacher Linking, Student Linking, Teacher Press, and Student Providing. After the feedback related to each of the constructs in feedback session two, there was improvement in each of the IQA toolkit rubric ratings for Accountable Talk.

It may be concluded that the feedback from lesson two had an influence on the instructional quality as it related to the sub-constructs for the Academic Rigor rubrics. The ways in which the influence was present was captured in the occurrences in the classroom during instruction. For lesson two, there was an instance of press from Miss S that contributed to the score. However, the recorded attempt to press did not appear to be deliberate on the part of the participant but occurred as the participant asked a brief follow-up to a question. Lesson three, however, recorded more instances of press, requiring more clarifications and justifications from students. In lesson one contributions to press came predominantly from Miss S revoicing her students' contributions. Lesson three, however, had more instances of press for students to provide conceptual responses and therefore lesson three seemed to have been influenced more from feedback session two than lesson two was influenced by feedback session one based on the differences in ratings from the tool. Since there were more instances of press from lesson three, students had more opportunities to provide conceptual responses, which they did. Feedback two indirectly affected the score for Student Providing since it influenced the participant's action during lesson three to press students for more conceptual responses.

The direct influences that were determined included the concept of linking as well where there was a noticeable difference on the part of the teacher to encourage students to connect their

ideas by comparing them and deciding if they agreed or not. Linking from lesson one was only recorded with revoicing. This changed dramatically by lesson three.

For Accountable talk, improvements were seen in the overall talk in the classroom. The teacher moved from dominating talk during the discussion to facilitating students' talk, allowing them to connect mathematical ideas with their peers. Students were also pressed to provide more conceptual responses and to justify claims that they made about the mathematics.

The participant seemed to have become more aware of her students and asking questions of them so they could demonstrate their thinking. She facilitated talk in her class much better after three iterative processes of teaching, feedback, and teaching informed by feedback.

Kluger and DeNisi (1996) suggested that there was an assumption from previous researchers that feedback would always lead to improved performance, and that assumption, which Kluger and DeNisi suggested was incorrect, influenced the development of the Feedback Intervention Theory (FIT). From the results of this study, and its connection to the FIT, the data suggested that the feedback given led to improved performance. Important to note is that the improvement in all the areas did not occur immediately following feedback session one. Based on the results, some areas that were discussed in the feedback session improved in the next observation, and others improved after more than one feedback session.

According to the results, the score for one sub-construct increased after feedback cycle one but decreased after the next even though the feedback was positive. This suggests that instruction can be inconsistent even after positive feedback is given. To focus on all aspects of the lesson takes much planning and follow-up execution. It could also mean that more feedback sessions are needed to help the participant maintain focus on all aspects of her lesson so that instructional quality can be improved.

Aspects of the results of this study can be linked to previous research. The influence of the use of feedback with the IQA tool has shown to improve the desired behaviors of the preservice teacher during instruction. Focused feedback on the process contributed to the reception of the preservice teacher to the feedback. Khachatryan (2015) found that process feedback led to more wanted behaviors among participants. The results of this study would be consistent with the results of Khachtryan since more positive behaviors were observed for the participant's instruction over the observation cycles.

While the methods of collecting and analyzing the data may be varied across researchers, there is agreement among some researchers on the effectiveness of feedback given to preservice teachers. In this study, there were improved scores for each of the sub-constructs from both Academic Rigor and Accountable Talk over the feedback cycles. This is consistent with Capizzi et al. (2010), who found that there were improvements on all indicators of the lesson component when preservice teachers were provided with feedback.

It makes sense that preservice teachers would value feedback from individuals they think are able to help them in their practice. Anderson and Radencich (2001) reported that preservice teachers found that feedback from their classroom teacher and their college supervisor was helpful. While the feedback given to the intern by the researcher influenced her instructional quality, it is also important to note that influences may also be attributed to other factors. Figure 17 demonstrates possible sources of influence on the overall instructional quality of the PST. The researcher's influence from the feedback sessions were apparent based on the results of the study. However, the PST had a university instructor who provided guidance on lesson planning, which could possibly contribute to some of the observed improvements on her instructional

quality. Additionally, the classroom teacher, with whom she spent most of her time undoubtedly provided direction to the PST on the content of the lessons, ways to implement the lesson and activities associated with the lesson. Another factor of influence is that of the peers of the PST. Interns speak to each other and discuss their experiences so the peers of the PST may provide an added avenue of influence, although the influence may be minimal based on the nature of the influence that was observed.

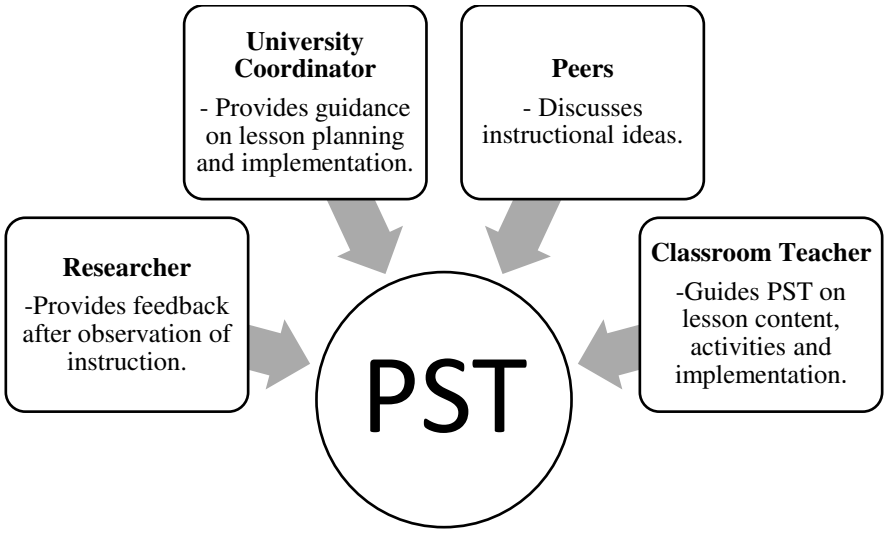


Image by author

Figure 17: Potential sources of influence for participant.

It is therefore important that individuals who are placed in positions to help by providing feedback are equipped with knowledge that improves the skill of the preservice teacher and those teachers are not propelled into the education system with a false sense of preparation. Focused feedback from reliable sources might make the difference between superficial lessons and lessons that are driven by the students themselves.

What seemed important in the participant's reflections and during the feedback sessions was that the participant began to notice aspects of her instruction that were more clearly connected to characteristics closely aligned with instructional quality in mathematics. This was an unexpected outcome of the application of the iterative process of teaching, feedback, and teaching based on feedback. Based on this outcome, it seemed important to determine what the participant noticed while observing the instruction of others.

The Role of Noticing

A key area of mathematics education is to equip teachers so that they are able to take informed actions during instruction. These actions are supported by the teachers being able to interpret the meaning of classroom interactions, analyze these interactions, and respond based on the analysis. The process of identifying actions and interpreting meaning is known as noticing (Van Es & Sherin, 2002). The emphases of the participant in her reflections changed in depth of focus over the three feedback cycles in which she was involved, from talking generally in her reflections to stating specific actions to support her claims. Noticing, as explored by Van Es and Sherin (2002), focused on identifying what is important in a teaching situation since the teacher is not able to respond to all the occurrences in the classroom. The researchers noted that identifying what is important is a key aspect of noticing.

Interaction with the use of the IQA toolkit, while in the feedback process, allowed for the participant to develop an understanding of what specific details were being observed during her observation. Evidence of knowing what to observe came from her reflections that were provided after each lesson. She talked about aspects of her lesson such as the task given to students, the types of questions asked during instruction, and her own development as a teacher to better

facilitate students during instruction. Knowing what was being observed in her lesson possibly made her more aware of how and when to respond to students. Teacher press requires teachers to ask follow-up questions to elicit more conceptual responses from students or for students to justify their claims. The participant's score for Teacher Press improved over the three feedback cycles because she was able to respond appropriately to students' answers by noticing there was the opportunity for students to extend their response to questions asked during the discussion.

The same concept of noticing may have allowed the participant to transfer that knowledge to the observation of the supervising teacher. Observing her supervising teacher was not about being able to score the IQA toolkit but being able to identify elements in instruction, as discussed during feedback, that are captured by the rubric. Taking an interpretive role for the purpose of understanding occurrences in the classroom, students' thought processes, or how teacher decisions influenced student thinking as opposed to critiquing, better explains the role of noticing (Van Es & Sherin, 2002). The participant, in her observation of the teacher, sought to play more of an interpretive role rather than seeking to critique the lesson, which was evident from her reflection. Explanations from the participant were connected to areas linked to the IQA toolkit based on the language she used in her reflection. Observing other teachers is an important role for preservice teachers during their training. The value of this experience may not be realized if the preservice teacher does not know where to place focus and what is important to observe (Star & Strictland, 2008). Berliner et al. (2008) suggest that the teacher's ability to notice is related to his or her classroom experience stating that the more experienced teacher notices more. However, Van Es and Sherin (2002) determined that the ability to notice is a skill that can be improved and may not only be evident in experienced teachers. The participant then,

although a preservice teacher, may have been able to develop noticing skills based on certain interactions during the iterative process of feedback cycles. In this study, her interaction with the IQA toolkit and its contents may have influenced her ability to notice occurrences both by the teacher and the students in the lesson she observed.

Limitations

For this study, there were some limitations that may be considered that prevented the study from being more robust. The most significant limitation to this research was the number of participants for the study. The initial number of participants targeted for the study was four so that there was the possibility of conducting cross-case analysis. However, enough participants did not enroll and two of the three that showed interest did not continue. Therefore leaving only one participant. This reduced the robust nature of the original research plan.

Another limitation that possibly contributed to the lack of interest for participants was that the word tool was used in the advertisement of the research. One potential participant said they thought they would have to learn to use the tool, which seemed too demanding and therefore the potential participant was not initially interested. This possibly happened because the PSTs were unfamiliar with the IQA toolkit. This could be corrected with a description of the use of the tool to include examples of the process as well.

During this research, the researcher played dual roles of the researcher and the observer of lessons. There is the possibility that this dual role caused some bias in the research that might not have been present if there were not these crucial roles held by the same person. There could be a research team who had different roles that could have reduced bias or conflict.

Implications for Future Research

As was previously stated, the IQA toolkit was traditionally used as an assessment instrument without providing feedback to the person being observed. Although there was more than one observation in most instances of published research, because validity of the use of the tool requires at least three observations for whole class settings (Wilhelm & Kim, 2015), there was no feedback provided between the observations. Based on the findings of this research, there are a few suggestions for future research in the area of feedback.

Firstly, from this research, with only three observations and two feedback sessions, there were observed differences in a number of the sub-constructs from Academic Rigor and Accountable Talk. Noticeable were areas that did not show up at all in lesson one, being observed on a large scale by lesson three, for example Teacher and Student Linking. The participant, being made aware that students may be given the opportunity to relate their ideas to each other, facilitated this, thereby providing greater interaction among students in the class. Future research would be better able to capture the influence of feedback because a larger research study may be able to consider other ideas that were not possible in this study. For example, analysis of time differences between when the feedback is provided and the next observed lesson to determine the level of instructional quality of participants over time, whether there is continuous improvement or a reduction. There may also be analysis to determine if certain parts of teacher's instruction remain constant or increase and what other aspects as measured by the IQA toolkit are neglected.

In future research as well, the number of participants could be increased to be able to conduct cross case analysis. This cross case analysis could determine differences, similarities,

and concerns among participants. With more participants, the most beneficial methods of conducting feedback using the IQA toolkit could be developed and a guide provided for how someone in authority such as an administrator observing mathematics classes may be able to provide more focused and specific feedback that can help teachers improve their instructional quality. This leads to the participants being extended from interns to teacher participants and feedback provided by trained observers to accurately score the IQA toolkit.

The IQA toolkit was used for this study to guide feedback on instruction. The tool drove the decision-making process of the feedback that was provided to the participant. The participant was able to have authentic experiences that reflected today's classroom expectations where the contributions of the students rather than the teacher become the focus and drive the lesson (Ittigson, 2002). The tool captured and encouraged authentic experiences students should have in mathematics and therefore influenced the feedback given to the PST, which in turn was reflected in instruction. While teacher preparation programs lack the rigor that is needed in effectively preparing students, the experience received in redesigning lessons than using pre-packaged ones are important for the success of the teacher (Darling-Hammond, 2010). The IQA toolkit was able to capture reform-oriented practices of classroom involvement by teachers and provided at-scale measurement results (Boston et al., 2015).

The study reflected the need to modify clinical processes to create the opportunity for PSTs to experience the learning that is important to shaping their expectations of classroom experiences. It also determined that the PST needed someone who was able to provide much more than a superficial observation but was also able to provide focused feedback on process rather than product (Khachatryan, 2015).

APPENDIX A: INSTRUCTIONAL QUALITY ASSESSMENT TOOLKIT

Instructional Quality Assessment Classroom Observation Tool

COVER PAGE – COMPLETE FOR EACH LESSON AND ATTACH TO FIELD NOTES, COPY OF INSTRUCTIONAL TASK, AND SCORE SHEET

Background Information

Date of observation: _____ **Observer:** _____

Start Time: _____ **End Time:** _____

District: _____ **School:** _____

Grade: _____ **Day 1 or Day 2** _____

Classroom Context

Total number of students in the classroom: _____
_____***Boys*** _____**Girls Sketch of seating arrangement(s):**

Mathematical Topic of the Lesson:

Field Notes (attach).

Academic Rigor

RUBRIC 1: Potential of the Task

Did the task have potential to engage students in rigorous thinking about challenging content?

4	**The task has the potential to engage students in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships, such as:** • Doing mathematics: using complex and non-algorithmic thinking (i.e., there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example); OR • Procedures with connections: applying a broad general procedure that remains closely connected to mathematical concepts. The task must explicitly prompt for evidence of students' reasoning and understanding. For example, the task **MAY** require students to: • solve a genuine, challenging problem for which students' reasoning is evident in their work on the task; • develop an explanation for why formulas or procedures work; • identify patterns and form and justify generalizations based on these patterns; • make conjectures and support conclusions with mathematical evidence; • make explicit connections between representations, strategies, or mathematical concepts and procedures. • follow a prescribed procedure in order to explain/illustrate a mathematical concept, process, or relationship.
3	**The task has the potential to engage students in complex thinking or in creating meaning for mathematical concepts, procedures, and/or relationships. However, the task does not warrant a "4" because:** • the task does not explicitly prompt for evidence of students' reasoning and understanding. • students may be asked to engage in doing mathematics or procedures with connections, but the underlying mathematics in the task is not appropriate for the specific group of students (i.e., too easy or too hard to promote engagement with high-level cognitive demands); • students may need to identify patterns but are not pressed for generalizations or justification; • students may be asked to use multiple strategies or representations but the task does not explicitly prompt students to develop connections between them; • students may be asked to make conjectures but are not asked to provide mathematical evidence or explanations to support conclusions
2	The potential of the task is limited to engaging students in using a procedure that is either specifically called for or its use is evident based on prior instruction, experience, or placement of the task. **There is little ambiguity about what needs to be done and how to do it.** The task does not require students to make connections to the concepts or meaning underlying the procedure being used. **Focus of the task appears to be on producing correct answers rather than developing mathematical understanding (e.g., applying a specific problem solving strategy, practicing a computational algorithm).** **OR** There is evidence that the mathematical content of the task is at least 2 grade-levels below the grade of the students in the class.

1	**The potential of the task is limited to engaging students in memorizing or reproducing facts, rules, formulae, or definitions. The task does not require students to make connections to the concepts or meaning that underlie the facts, rules, formulae, or definitions being memorized or reproduced.**
0	**The task requires no mathematical activity.**
N/A	Students did not engage in a task.

ATTACH OR DESCRIBE THE TASK.

RUBRIC 2: Implementation of the Task

At what level did the teacher guide students to engage with the task in implementation?

4	**Students engaged in exploring and understanding the nature of mathematical concepts, procedures, and/or relationships, such as:** • Doing mathematics: using complex and non-algorithmic thinking (i.e., there is not a predictable, well-rehearsed approach or pathway explicitly suggested by the task, task instructions, or a worked-out example); OR • Procedures with connections: applying a broad general procedure that remains closely connected to mathematical concepts. There is explicit evidence of students' reasoning and understanding. For example, students may have: • solved a genuine, challenging problem for which students' reasoning is evident in their work on the task; • developed an explanation for why formulas or procedures work; • identified patterns, formed and justified generalizations based on these patterns; • made conjectures and supported conclusions with mathematical evidence; • made explicit connections between representations, strategies, or mathematical concepts and procedures. • followed a prescribed procedure in order to explain/illustrate a mathematical concept, process, or relationship.
3	**Students engaged in complex thinking or in creating meaning for mathematical concepts, procedures, and/or relationships. However, the implementation does not warrant a "4" because:** • there is no explicit evidence of students' reasoning and understanding. • students engaged in doing mathematics or procedures with connections, but the underlying mathematics in the task was not appropriate for the specific group of students (i.e., too easy or too hard to sustain engagement with high-level cognitive demands); • students identified patterns but did not make generalizations; • students used multiple strategies or representations but connections between different strategies/representations were not explicitly evident; • students made conjectures but did not provide mathematical evidence or explanations to support conclusions
2	Students engaged in using a procedure that was either specifically called for or its use was evident based on prior instruction, experience, or placement of the task. **There was little ambiguity about what needed to be done and how to do it.** Students did not make connections to the concepts or meaning underlying the procedure being used. **Focus of the implementation appears to be on producing correct answers rather than developing mathematical understanding (e.g., applying a specific problem solving strategy, practicing a computational algorithm).** OR There is evidence that the mathematical content of the task is at least 2 grade-levels below the grade of the students in the class.
1	**Students engage in memorizing or reproducing facts, rules, formulae, or definitions. Students do not make connections to the concepts or meaning that underlie the facts, rules, formulae, or definitions being memorized or reproduced.**
0	The students did not engage in mathematical activity.

N/A	**The students did not engage with a mathematical task.**

RUBRIC 3: Student Discussion Following Task

To what extent did students show their work and explain their thinking about the important mathematical content?

4	Students show/describe written work for solving a task and/or engage in a discussion of the important mathematical ideas in the task. During the discussion, students provide complete and thorough explanations of why their strategy, idea, or procedure is valid; students explain why their strategy works and/or is appropriate for the problem; students make connections to the underlying mathematical ideas (e.g., "I divided because we needed equal groups"). OR Students show/discuss more than one strategy or representation for solving the task, provide explanations of why the different strategies/representations were used to solve the task, *and/or make connections between strategies or representations. [Thorough presentation and discussion across strategies or representation.]*
3	Students show/describe written work for solving a task and/or engage in a discussion of the important mathematical ideas in the task. During the discussion, students provide explanations of why their strategy, idea, or procedure is valid and/or students begin to make connections BUT the explanations and connections are not complete and thorough (e.g., student responses often require extended press from the teacher, are incomplete, lack precision, or fall short making explicit connections). OR Students show/discuss more than one strategy or representation for solving the task, and provide explanations of how the different strategies/representations were used to solve the task *but do not make connections between different strategies or representations. [Thorough presentation and/or discussion of individual strategies or representations, but no talk or questioning across different strategies)]*
2	Students show/describe written work for solving the task (e.g., the steps for a multiplication problem, finding an average, or solving an equation; what they did first, second, etc) but do not engage in a discussion of why their strategies, procedures, or mathematical ideas work; *do not make connection to mathematical concepts. [Procedural explanations only]* OR Students show/discuss only one strategy or representation for solving the task, OR Students make presentations of their work with no questioning or prompting from the teacher (to the presenters or to the class) to explain the mathematical work, make connections, etc. [*Presentations with no discussion*]
1	Students provide brief or one-word answers (e.g., fill in blanks); OR **Student's responses are non-mathematical.**
0	There was no discussion of the task.

N/A	Reason:

Rigor of Teachers' Questions

Rubric AR-Q: Questioning	
4	The teacher **consistently** asks academically relevant questions that provide opportunities for students to elaborate and explain their mathematical work and thinking (probing, generating discussion), identify and describe the important mathematical ideas in the lesson, or make connections between ideas, representations, or strategies (exploring mathematical meanings and relationships).
3	At least 2 times during the lesson, the teacher asks academically relevant questions (probing, generating discussion, exploring mathematical meanings and relationships).
2	There are one or more superficial, trivial, or formulaic efforts to ask academically relevant questions probing, generating discussion, exploring mathematical meanings and relationships) (i.e., every student is asked the same question or set of questions) or to ask students to explain their reasoning; OR only one (1) strong effort is made to ask academically relevant questions.
1	The teacher asks procedural or factual questions that elicit mathematical facts or procedure or require brief, single word responses.
0	The teacher did not ask questions during the lesson, or the teacher's questions were not relevant to the mathematics in the lesson.
N/A	Reason:

Rubric 1: Participation	
4	Over 75% of the students participated throughout the discussion.
3	50-75% of the students participated in the discussion.
2	25-50% of the students participated in the discussion.
1	Less than 25% of the students participated in the discussion.
0	None of the students participated in the discussion.
N/A	Reason:

_____Number of students in class

_Number of students who participated

Teacher's Linking Contributions

Does the teacher support students in connecting ideas and positions to build coherence in the discussion?

Rubric 2: Teacher's Linking	
4	The teacher consistently (at least 3 times) connects (or provides opportunities for students to connect) speakers' contributions to each other and shows (or provides opportunities for students to show) how ideas/positions shared during the discussion relate to each other.
3	At least twice during the lesson the teacher connects (or provides opportunities for students to connect) speakers' contributions to each other and shows (or provides opportunities for students to show) how ideas/positions relate to each other.
2	At one or more points during the discussion, the teacher links speakers' contributions to each other, but does not show how ideas/positions relate to each other (weak links). No follow-up questions are asked after speakers' contributions OR teacher revoices or recaps only, but does not show how ideas/positions relate to each other OR only one strong effort is made to connect speakers' contributions to each other (1 strong link).
1	Teacher does not make any effort to link or revoice speakers' contributions.
0	No class discussion OR Class discussion was not related to mathematics.
N/A	Reason:

Students' Linking Contributions Do student's contributions link to and build on each other?

Rubric 3: Students' Linking	
4	The students consistently connect their contributions to each other and show how ideas/positions shared during the discussion relate to each other. (e.g. I agree with Jay because…")
3	At least twice during the lesson the students to connect their contributions to each other and show how ideas/positions shared during the discussion relate to each other. (e.g. I agree with Jay because…")
2	At one or more points during the discussion, the students link students' contributions to each other, but do not show how ideas/positions relate to each other. (e.g., "I disagree with Ana.") OR only one strong effort is made to connect their contributions with each other.
1	Students do not make any effort to link or revoice students' contributions.
0	No class discussion OR Class discussion was not related to mathematics.

N/A	Reason:

How effectively did the lesson-talk build Accountability to Knowledge and Rigorous Thinking?

Asking: Were students pressed to support their contributions with evidence and/or reasoning?

Rubric 4: Asking (Teachers' Press)	
4	The teacher consistently (almost always) asks students to provide evidence for their contributions (i.e., press for conceptual explanations) or to explain their reasoning. (There are few, if any instances of missed press, where the teacher needed to press and did not.)
3	Once or twice during the lesson the teacher asks students to provide evidence for their contributions (i.e., press for conceptual explanations) or to explain their reasoning. (The teacher sometimes presses for explanations, but there are instances of missed press.)
2	Most of the press is for computational or procedural explanations or memorized knowledge OR There are one or more superficial, trivial efforts, or formulaic efforts to ask students to provide evidence for their contributions or to explain their reasoning (i.e., asking everyone, "How did you get that?").
1	There are no efforts to ask students to provide evidence for their contributions, AND there are no efforts to ask students to explain their thinking.
0	Class discussion was not related to mathematics OR No class discussion
N/A	Reason:

Providing: Did students support their contributions with evidence and/or reasoning? (This evidence must be appropriate to the content area—i.e., evidence from the text; citing an example, referring to prior classroom experience.)

Rubric 5: Providing (Students' Responses)	
4	Students consistently provide evidence for their claims, OR students explain their thinking using reasoning in ways appropriate to the discipline (i.e. conceptual explanations).

APPENDIX B: INSTITUTIONAL REVIEW BOARD PERMISSION LETTER

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Makini Ashaki Campbell Sutherland**

Date: **March 27, 2017**

Dear Researcher:

On 03/27/2017 the IRB approved the following modifications to human participant research until 01/26/2018 inclusive:

Type of Review:	Submission Response for IRB Addendum and Modification Request Form
Modification Type:	Additional Study Locations
Project Title:	A Collective case study on the influence of feedback on the instructional quality among preservice elementary teachers in mathematics: The instructional quality assessment tool as a resource
Investigator:	Makini Ashaki Campbell Sutherland
IRB Number:	SBE-16-12820
Funding Agency:	
Grant Title:	
Research ID:	n/a

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 01/26/2018, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Signature applied by Renea C Carver on 03/27/2017 11:41:30 AM EDT

IRB Coordinator

APPENDIX C: DISTRICT IN CENTRAL FLORIDA SCHOOL'S PERMISSION

[redacted]

February 12, 2017

Makini Sutherland

[redacted]

Dear Ms. Sutherland

I am in receipt of the proposal and supplemental information that you submitted for permission to conduct research in th[redacted] Thank you for very clearly delineating the required components of the research request. After a review of these documents, it has been determined that you are granted permission to conduct the study described herein.

Your project, *A Collective Case Study on the Influence of Feedback Among Preservice Elementary Teachers in Mathematics: The Instructional Quality Assessment Tool (IQA) As a Resource*, is of interest to the District. Your first order of business is to contact the principal of each of your targeted schools. (Please see the list below.) You indicated in your Research Permission Request that you will select four (4) of the six (6) schools for your study.

Please be advised that securing permission from th[redacted] Preservice Teachers' Supervisor(s) is your responsibility. Please be sure you have the appropriate approval prior to contacting the principals.

We look forward to receiving a copy of your results. Best of luck!

Respectfully,

[redacted]

Deputy Superintendent, Instructional Excellence and Equity

cc. [redacted]

APPENDIX D: REFLECTION PROTOCOL

How do you feel about the delivery of the lesson observed?

- Were all your objectives achieved?
  - If no what contributed to this?
  - If yes what evidence can you provide?
- Is there any aspect of your lesson that did not go as planned?
  - If yes, how did you handle the situation?
  - Was your handling effective?
  - Why do you think so?

Consider the task(s) that was given

- Do you think you provided a rich task for your students?
  - What evidence do you have to support your answer?

Consider the general implementation of your lesson

- To what extent do you think the lesson implementation was effective?
  - What evidence do you have to support your response?

Consider your class discussion

- Did you have a whole class discussion?
  - If no, why not?
  - If yes, do you think it was effective?
  - What evidence supports your response?

Consider questions you asked of students

- Did you ask academically relevant questions?
  - If yes, provide examples of these
  - State whether these occurred during whole class or individual work

APPENDIX E: PERMISSION TO USE IQA TOOLKIT

Re: Permission to use IQA rubric descriptions

Makini Sutherland

Mon 10/9/2017 5:28 PM

Sent Items

To: Dr. Melissa Boston <bostonm@duq.edu>;

Thank you Dr. Boston

Makini Sutherland
Doctoral Candidate, Mathematics Education
AACTE Holmes Scholar
University of Central Florida

From: Dr. Melissa Boston <bostonm@duq.edu>
Sent: Monday, October 9, 2017 3:40:35 PM
To: Makini Sutherland
Cc: Juli Dixon, Ph.D.
Subject: RE: Permission to use IQA rubric descriptions

Hi Makini—you have my permission to include the rubrics. After the figure name, include "Boston (2012) Used with permission of the author."

Best of luck on your dissertation!
Melissa

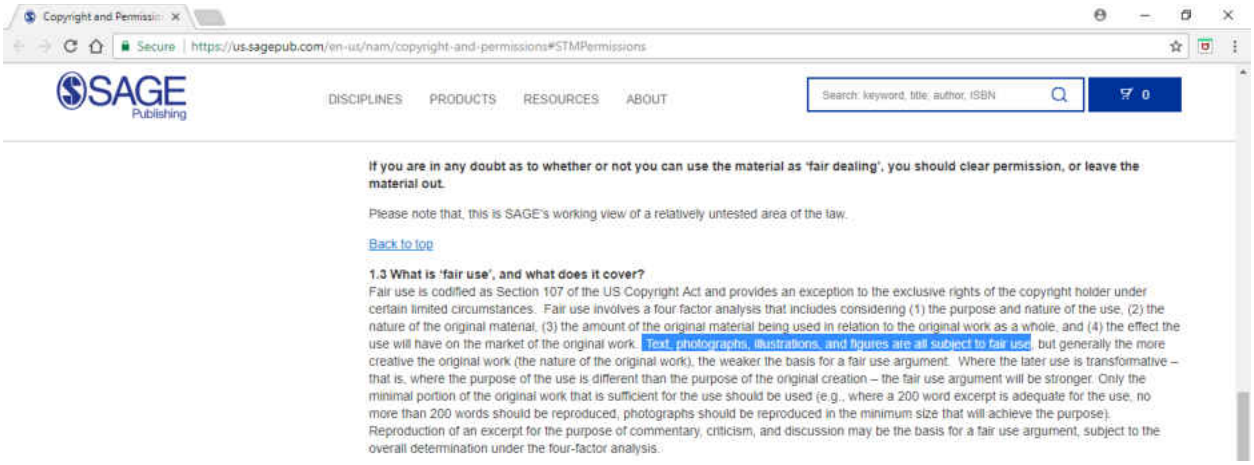
From: Makini Sutherland [mailto:makini.sutherland@Knights.ucf.edu]
Sent: Monday, October 9, 2017 2:32 PM
To: Dr. Melissa Boston <bostonm@duq.edu>
Cc: Juli Dixon, Ph.D. <juli.dixon@ucf.edu>
Subject: Permission to use IQA rubric descriptions

Good day Dr. Boston,

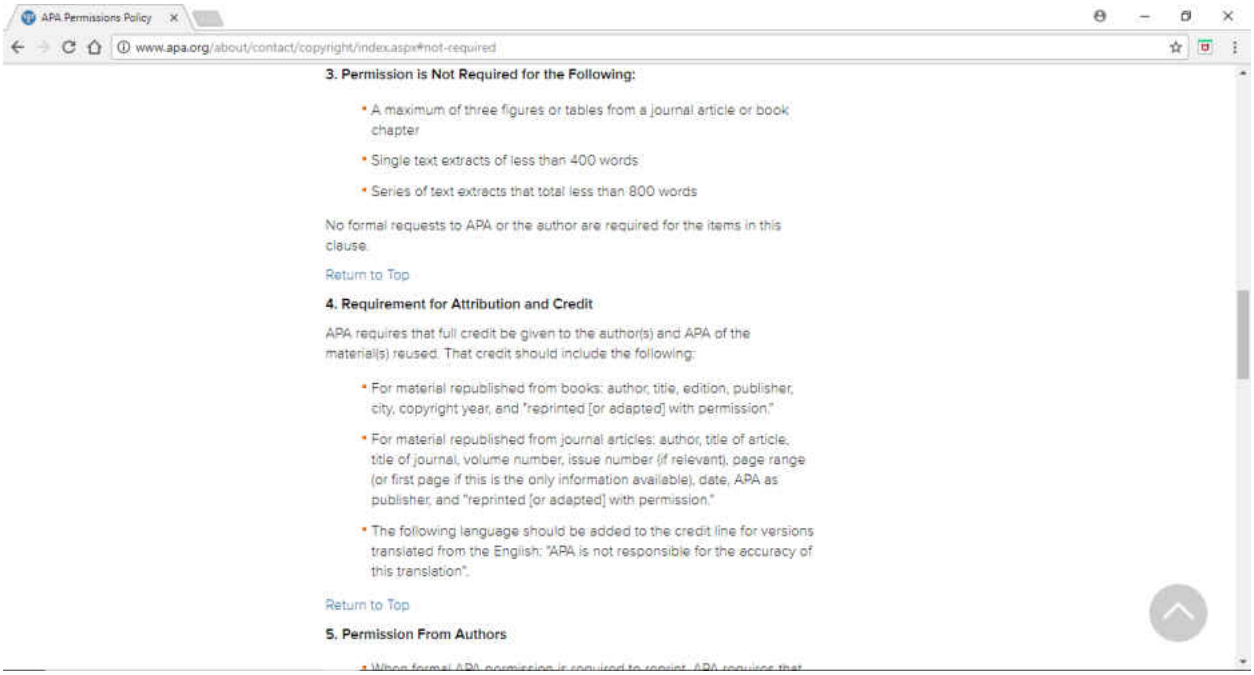
I am in the process of writing up my dissertation and am explaining differences in the score levels for the Potential of the Task, ,Implementation of the Task, Discussion and Participation. To do this I want to insert the descriptions related to some of the scores. Dr. Dixon told me I need permission to do this so I am requesting permission to include the inserts
in my write-up. Please let me know if you need more descriptions of what I want to do.

Regards
Makini Sutherland
Doctoral Candidate, Mathematics Education
AACTE Holmes Scholar
University of Central Florida

APPENDIX F: PERMISSION TO USE SAGE PUBLICATIONS DIAGRAM


Copyright and Permissio
Secure | https://us.sagepub.com/en-us/nam/copyright-and-permissions#STMPermissions
SAGE Publishing
DISCIPLINES
PRODUCTS
RESOURCES
ABOUT
Search: keyword, title, author, ISBN
0
If you are in any doubt as to whether or not you can use the material as 'fair dealing', you should clear permission, or leave the material out.
Please note that, this is SAGE's working view of a relatively untested area of the law.
Back to top
1.3 What is 'fair use', and what does it cover?
Fair use is codified as Section 107 of the US Copyright Act and provides an exception to the exclusive rights of the copyright holder under certain limited circumstances. Fair use involves a four factor analysis that includes considering (1) the purpose and nature of the use, (2) the nature of the original material, (3) the amount of the original material being used in relation to the original work as a whole, and (4) the effect the use will have on the market of the original work. Text, photographs, illustrations, and figures are all subject to fair use, but generally the more creative the original work (the nature of the original work), the weaker the basis for a fair use argument. Where the later use is transformative – that is, where the purpose of the use is different than the purpose of the original creation – the fair use argument will be stronger. Only the minimal portion of the original work that is sufficient for the use should be used (e.g., where a 200 word excerpt is adequate for the use, no more than 200 words should be reproduced, photographs should be reproduced in the minimum size that will achieve the purpose). Reproduction of an excerpt for the purpose of commentary, criticism, and discussion may be the basis for a fair use argument, subject to the overall determination under the four-factor analysis.

APPENDIX G: PERMISSION TO USE APA PUBLICATIONS DIAGRAM


APA Permissions Policy
www.apa.org/about/contact/copyright/index.aspx#not-required
3. Permission is Not Required for the Following:
A maximum of three figures or tables from a journal article or book chapter
Single text extracts of less than 400 words
Series of text extracts that total less than 800 words
No formal requests to APA or the author are required for the items in this clause.
Return to Top
4. Requirement for Attribution and Credit
APA requires that full credit be given to the author(s) and APA of the material(s) reused. That credit should include the following:
For material republished from books: author, title, edition, publisher, city, copyright year, and "reprinted [or adapted] with permission."
For material republished from journal articles: author, title of article, title of journal, volume number, issue number (if relevant), page range (or first page if this is the only information available), date, APA as publisher, and "reprinted [or adapted] with permission."
The following language should be added to the credit line for versions translated from the English: "APA is not responsible for the accuracy of this translation".
Return to Top
5. Permission From Authors

REFERENCES

Anderson, N.A., & Radencich, M.C. (2001). The value of feedback in an early field experience: Peer, teacher, and supervisor coaching. *Action in Teacher Education*, 23(3), 66-74

Ball, D. L. (2003). Mathematical proficiency for all students. Retrieved from http://fs1.bib.tiera.ru/content/ShiZ/math/other/Mathematical%20Proficiency%20For%20All%20Students%20-%20RAND%20(2003).pdf

Ball, D., Lubienski, S. T., & Spangler-Mewborn, D. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of Research on Teaching* (4th ed.). New York: Macmillan.

Ball, D. L. (1990). Prospective elementary and secondary teachers' understanding of division. *Journal for Research in Mathematics Education, 21*(2), 132-144.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389-407.

Benson, F. (2014). A decade of evolution in the practice teaching component of a Canadian teacher education program: What drove change, what insights were gleaned, and what challenges lie ahead. *Global Education Review, 1*(3), 7-62.

Bieda, K. N., Sela, H., & Chazan, D. (2015). "You are learning well my dear": Shifts in novice teachers' talk about teaching during their internship. *Journal of Teacher Education, 66*(2), 150-169.

Boaler, J. & Brodie, K. (2004). *The importance, nature and impact of teacher questions.* Paper presented at the North American Chapter of the International Group for the Psychology of Mathematics Education Toronto, Ontario.

Boaler, J. & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside school. *Teacher College Record*, 110, 608-645.

Boston, M. (2012). Assessing instructional quality in mathematics. *The Elementary School Journal, 113*(1), 76-104.

Boston, M., Bostic, J., Lesseig, K., & Sherman, M. (2015). A comparison of mathematics observation protocols. *Mathematics Teacher Educator, 3*(2), 154-175.

Boston, M. & Smith, M. (2009). Transforming secondary mathematics teaching: Increasing cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education, 40*(2), 119-156.

Boston, M. & Wilhelm, A. G. (2015). Middle school mathematics instruction in instructionally-focused urban districts. *Urban Education*. doi:10.1177/0042085915574528

Boston, M. & Wolf, M. K. (2006). Assessing academic rigor in mathematics instruction: The development of the Instructional Quality Assessment toolkit. Los Angeles: University of California.

Capizzi, A. M., Wehby, J. H., & Sandmel, K. N. (2010). Enhancing mentoring of teacher candidates through consultative feedback and self-evaluation of instructional delivery. *Teacher Education and Special Education*, 33, 191212.

Carnoy, M. & Rothstein, R. (2013). What do international tests really show about U.S. student performance? *Economic Policy Institute*.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (2015). *Children's mathematics: Cognitively guided instruction*.

Carver, C. S. & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality, social, clinical and health psychology. *Psychological Bulletin, 92*(1), 705-717.

Cochran, K. F., King, R. A., & DeRuiter, J. (1991). *Pedagogical content knowledge: A tentative model for teacher preparation*. Paper presented at the American Education Research Association, Chicago, IL.

Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among the five approaches* (3rd ed.): SAGE Publications Inc.

Crommelinck, M. & Anseel, F. (2013). Understanding and encouraging feedback-seeking behaviour: A literature review. *Medical education*. 47. 232-41. 10.1111/medu.12075.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). Alexandria VA: ASCD.

Darling-Hammond, L. (2010). Teacher education and the American future. *Journal of Teacher Education, 61*(1-2), 35-47.

Darling-Hammond, L. (2014). Strengthening clinical preparation: The holy grail of teacher education. *Peabody Journal of Education, 89*(4), 547-561.

Florida Department of Education, (2016). Annual legislative report on teacher evaluation. Retrieved from fldoe.org/core/fileparse.php/7503/urlt/1314AnnualLegisReportTeacherEval.pdf.

Gall, M., Gall, J., & Borg, W. (2006). *Educational Research: An Introduction* (8th ed.). Boston: MA: Allyn & Bacon.

Graybeal, C. D. (2013). Learning to look for the standards for mathematical practice. *SRATE Journal, 22*(2), 8-13.

Hancock, D. R. & Algozzine, B. (2006). *Doing case study research*. New York: Teachers College Press.

Hiebert, J. & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American educational research journal, 30*(2), 393-425.

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An Exploratory Study, Cognition and Instruction. Cognition and Instruction, 26(4), 430-511

Hoyt, K. & Terantino, J. (2015). Rethinking field observations: Strengthening teacher education through INFORM. *Action in Teacher Education, 37*(3), 209-222.

Hull, T., Miles, R. H., & Balka, D. S. (2014). Defining and instituting rigor *Realizing Rigor in the mathematics classroom* (pp. 2-12). Thousand Oaks: Corwin Press.

Imm, K., & Stylianou, D. A. (2012). Talking mathematically: An analysis of discourse communities. *The Journal of Mathematical Behavior, 31*, 130-148.

Isiksal, M. & Cakiroglu, E. (2011). The nature of prospective mathematics teachers' pedagogical content knowledge: The case of multiplication of fractions. *Journal of Mathematics Teacher Education, 14*, 213-230.

Ittigson, R. (2002). Helping students become mathematically powerful. *Teaching Children Mathematics, 9*(2), 91-95.

Junker, B. W., Crosson, A., Boston, M., Levison, A., Matsumura, L. C., Resnick, L., & Wolf, M. K. (2006). Beyond summative evaluation: The Instructional Quality Assessment as a professional development tool. Los Angeles: California.

Khachatryan, E. (2015). Feedback on teaching from observations of teaching: What do administrators say and what do teachers think about it? *SAGE Publications, 99*(2), 164-188.

Klein, A. (2016). Rethink teacher-evaluation systems if they're not working, John King says. Retrieved from http://blogs.edweek.org/edweek/campaign-k-12/2016/01/john_king_if_teacher_evaluatio.html.

Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284.

Latham, G. & Locke, E. (2007). New developments in and directions for goal setting research. *European Psychologist, 12*(4), 290-300.

Locke, E. & Latham, G. (2002). Building a practically useful theory of goal setting and task motivation: A 35 year old odyssey. *Psychological Bulletin, 92*(1), 111-135.

Locke, E., Shaw, K. N., Saari, L. M., & Latham, G. (1981). Goal setting and task performance: 1960-1980. *Psychological Bulletin, 90*(1), 125-152.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum.

Maher, N., & Muir, T. (2013). "I know you have to put down a zero, but i'm not sure why": Exploring the link between pre-service teachers' content and pedagogical content knowledge. *Mathematics Teacher Education and Development, 15*(1), 72-87.

Martin, C. K., & Mertl, V. (2014). Teaching to the core: Practitioner perspectives about the intersection of teacher evluation using the Danielson Framework for Teaching and Common Core State Standards.

Marzano, R. J. (2011). The Marzano Teacher Evaluation Model. *Marzano Research Laboratory*.

Matsumura, L., Garnier, H., Slater, S., & Boston, M. (2008). Towards measuring instructional interactions "at scale". *Educational Assessment, 13*, 267-300.

Matsumura, L. C., S. C., S., Junker, B., Peterson, M., Boston, M., Steele, M., & Resnick, L. (2006). Measuring reading comprehension and mathematics instruction in urban middle schools: A pilot study of the Instructional Quality Assessment. Los Angeles: University of California, National

Matsumura, L. C. (2000). *Creating high-quality classroom assignments*. Lanham, Md: Scarecrow Education.

McDonnough, J. & Matkins, J. J. (2010). The role of field experience in elementary preservice teachers' self-efficacy and ability to connect research to practice. *School science and mathematics, 110*(1), 13-23.

Meade, E. (1991). Reshaping the clinical phase of teacher preparation. *Phi Delta Kappan, 72*(9), 666-669.

Mewborn, D. (2001). Teachers content knowledge, teacher education, and their effects on the preparation of elementary teachers in the United States. *Mathematics Education Research Journal, 3*, 28-36.

Munter, C. (2014). Developing Visions of High-Quality Mathematics Instruction *Journal for research in mathematics education, 45*(4), 584-635.

Nathan, M. J. & Knuth, E. J. (2003). A study of whole class mathematical discourse and teacher change. *Cognition and Instruction, 21*, 175-207.

National Council for Teachers of Mathematics (2014) Mathematics. *Principles to Actions: Ensuring mathematical success for all*. Reston, VA.

National Research Council, (2001). *Adding it up: Helping children learn mathematics* (J. Kilpatrick, J. Swafford & B. Findell Eds.): National Council for teachers of mathematics

O'Shea, L. J., Hoover, N. L., & Carroll, R. G. (1988). Effective intern conferencing. *Journal of Teacher Education* (March-April).

Quint, J. C., Akey, T. M., Rappaport, S., & Williner, C. J. (2007). *Instructional leadership, teaching quality, and student achievement: Suggestive evidence from three urban school districts*. New York: MDRC.

Rhoads, K., Radu, I., & Weber, K. (2011). The teacher internship experiences of prospective high school mathematics teachers. *International Journal of Science and Mathematics Education, 9*, 999-1022.

Roegman, R., Goodwin, A. L., Reed, R., & II, R. M. S.-M. (2014). Unpacking the data: An analysis of the use of Danielson's (2007) Framework for Professional Practice in a teaching residency program. *Educational Assessment, Evaluation and Accountability, 28*, 111-137.

Roy, G. (2008). *Prospective teachers' development of whole number concepts and operations during a classroom teaching experiment*. Unpublished Doctoral Dissertation.

Ryan, J. & Williams, J. (2007). Mathsmaps for diagnostic assessment with pre-service teachers: stories of mathematical knowledge. *Research in Mathematics Education, 9*(1), 95-109.

Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2002). Measuring Reform Practices in Science and Mathematics Classrooms: The Reformed Teaching Observation Protocol. *School Science and Mathematics*, 102(6), 245-253.

Shulman, L. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher, 4*(14).

Stake, R. E. (2006). *Multiple case study analysis*. New York: The Guilford Press.

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move from show and tell. *Mathematical Thinking and learning, 10*(4), 313-340.

Stein, M. K. & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*(1), 50-80.

Stein, M. K., Remillard, J., Smith, M. S. (2007). How curriculum influences student learning. In F. Lester (Ed.), Second Handbook of Research on Mathematics Teaching and Learning (pp. 319-370). Charlotte, NC: Information Age Publishing.

Stein, M. K. & Smith, M. S. (1998). Mathematical Tasks as a Framework for Reflection: From Research to Practice. *Mathematics Teaching in the Middle School, 3*, 268-275.

Steinberg, M. & Sartain, L. (2015). Does better observation make better teachers? *Education Next*, 70-77.

Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61*(5), 12–17.

The Every Student Succeeds (ESSA) Act. (2015). (Pub. L. No. 114-95).

The nations report card 2015. Retrieved from https://www.nationsreportcard.gov/reading_math_2015/#?grade=4

The No Child Left Behind (NCLB) Act. (2001). (Pub. L. No. 107-110).

Wilhelm, A. G. & Kim, S. (2015). Generalizing from observations of mathematics teachers' instructional practice using the Instructional Quality Assessment (pp. 270-279): National Council of Teachers of Mathematics.

Wilson, J. D. (1996). An evaluation of the field experiences of the innovative model for the preparation of elementary teachers for science, mathematics and technology. *Journal of Teacher Education, 47*(1), 53-59.

Zeichner, K. (2010). Rethinking the connections between campus courses and field experiences in college- and university-based teacher education. *Journal of Teacher Education, 61*(1-2), 89-99.