13th International Congress on Mathematical Education
Hamburg, 24-31 July 2016

MEASURING MATHEMATICS CLASSROOM INTERACTIONS: AN OBSERVATION PROTOCOL REINFORCING THE DEVELOPMENT OF CONCEPTUAL UNDERSTANDING

Dr. Jeremy Zelkowski, Dr. Jim Gleason, Dr. Stefanie Livers

The University of Alabama

This study measured the degree to which actions of teachers and students in mathematics classrooms align with practices recommended by national organizations and initiatives, through the lenses of teaching mathematics for conceptual understanding within a community of learners, examining teacher facilitation, and student engagement. This theoretical model used the creation and validation of an observational instrument designed for the K-16 mathematics classroom, the Mathematics Classroom Observation Protocol for Practices (MCOP^2).Consideration was given to both direct and dialogic instruction. Instrument validation involved feedback from over 125 external review experts in mathematics education for content validity; inter-rater & internal reliability analyses; and structure analyses via scree plot analyses & exploratory factor analysis. A two-factor structure instrument was determined to support the measurement of teacher as a facilitator and student engagement in learning mathematics.

THE MATHEMATICS CLASSROOM PRACTICES

Over three decades of research has yielded relatively consistent findings with respect to the mathematical practices for which students ought to be engaged (Ball, Ferrini-Mundy, Kilpatrick, Milgram, Schmid, & Schaar, 2005) *to learn mathematics*. However, that is not to say there has not been differing views regarding the pedagogies teachers of mathematics should employ in the classroom (Munter, Stein, Smith, 2015). Ball et al. (2005) cultivated relative harmony regarding mathematical pedagogies that resulted in three vital components: (1) numerical fluency as being critical, (2) precise communication with purposeful reasoning, and (3) problem solving skills. Munter, Stein, and Smith (2015) built upon this harmony by convening a group of 26 leaders in the mathematics education community to deeply examine the similarities and differences of these three vital components within the differing pedagogies of direct instruction and dialogic instruction (also referred to as reformed teaching). This group of *mathematics education* leaders common aspirations regarding the teaching of mathematics resulted with a list that included: conceptual understanding, procedural fluency, rigorous mathematical tasks, student reasoning, and opportunities for student practice. Moreover, the group spotlighted differences regarding the role of (1) mathematical discourse, (2) student group work, (3) the description of tasks, (4) feedback, (5) creativity, (6) analyzing student work, (7) mathematical definitions, and (8) representations.

The mathematics education community identifies dialogic instruction as the preferred means of teaching and learning during the standards-based teaching era launched in 1989 (NCTM, 1989, 2000; NGACBP, CCSSO, 2010). Historically, the RTOP (Reformed Teaching Observation Protocol) (Sawada et al., 2010) was used as the main tool for evaluation and research endeavors in the field of mathematics and science classrooms regarding variation of teaching practices towards a dialogical direction of how lessons are enacted in the classroom. Currently, the assessment of teaching effectiveness often focuses on student standardized test scores from only one point in time during the school year as a primary means of measure. Observation protocols, student work samples, video case analyses, students' prior and end-of-year test scores, teacher portfolios, as well as a myriad of other data sources can provide teachers, evaluators, and researchers many differing viewpoint to measure teaching effectiveness. The mathematics education community in the United States has moved into the time of Common Core State Standards implementation and thus requiring a change in classroom practices, but change takes time and professional development.

Based on the complexity of such mathematical pedagogies, attempts to research and measure classroom practices becomes highly complex, but one that must include aspects of both direct and dialogic instructional approaches. The desideratum for mathematics classroom observation protocols for the practices of classroom teachers and students has reached a point of needed access beyond just the mathematics education research community. For the last 30-50 years, the field has generated strong empirical research regarding teaching practices, student achievement, and connections between both practices and achievement. However, measurement of classroom practices of teachers and students has really only been accessible to researchers. Our team of researchers at the University of Alabama developed a classroom observation protocol with aims to provide access to school-level personnel *and* the research community with the ability to measure practices for classroom, school, and district use, in addition to larger scale studies.

THEORETICAL FRAMEWORK

Mathematics Classroom Interactions for Conceptual Understanding

Conceptual understanding in mathematics has a long history, though called somewhat different in terms of research and practice for nearly a century (Brownell, 1935; Hiebert & Carpenter, 1992; Hiebert & Grouws, 2007; Skemp, 1971, 1979). When students lack conceptual understanding and have only basic procedural knowledge of mathematics, they fail at the opportunities to move beyond any mathematics other than procedural knowledge and repetition. Therefore, it is imperative that both teacher *and* student engage in productive discourse, tasks and reasoning (Barker et al., 2004; NCTM, 2006, 2011; Stein, Engle, Smith, & Hughes, 2008; Stein, Smith, Henningsen, & Silver, 2009). We shadow the definition of Hiebert and Grouws (2007) that teaching is "classroom interactions among teachers and students around content directed toward facilitating students' achievement of learning goals" (p. 377). Hence, this classroom model aligns well with the Instruction as Interaction framework of Cohen, Raudenbush, and Ball (2003) that has a strong eye on the discourse and interactions between the teacher, the mathematical content of the lesson, and the students' participation (see Figure 1). The framework reveals the multifaceted relationship within this triad, highlighting how these paradigms create or portray the classroom environment. Our attention on classroom interactions includes information about the atmosphere for teaching and learning, such as quantity and quality of teacher and student interactions with each other and with the mathematical content. It does not include other worthy information related to the learning setting such as school administration, socio-economic factors, or high-stakes assessments.

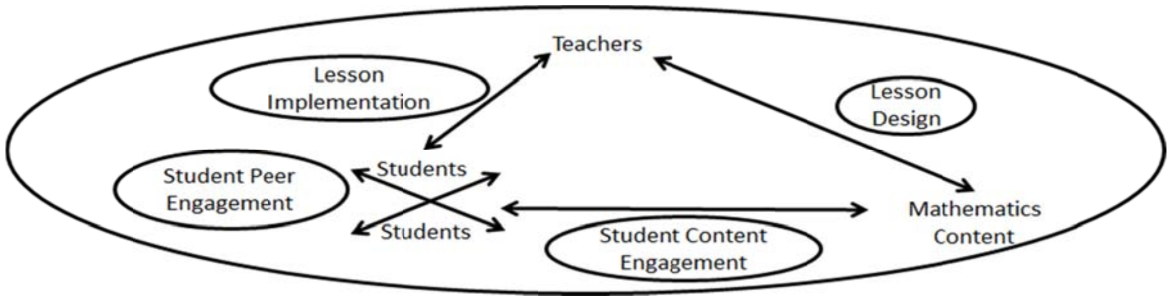

Figure 1. Instruction as Interaction Model (Based upon Cohen, Raudenbush, & Ball, 2003).

Within this framework, we posit the mathematics classroom as a *community of learners* (Rogoff, Matusov, & White, 1996) instructional model with the authority and responsibility shared between teachers and students. The role of the teacher as a facilitator is one who provides structure for the lesson, targets students' *zone of proximal development* (Vygotsky, 1978), gives guidance in the problem solving process (Polya, 1957), and promotes positive productive discourse norms within the classroom (Cobb, Wood, & Yackel, 1993; Nathan & Knuth, 2003). The role of the student engagement includes actively participating in the lesson content while persisting through the problem solving process (Stein et al, 2009) and engaging in productive discourse (Stein et al, 2008) as essential for making the connections necessary for the conceptual understanding of mathematics during the learning process.

Current efforts for improvements and research in mathematics teaching and learning aims for classrooms with teachers who can facilitate actions (implementation) that include interactions between students (peer engagement), implement lessons with high quality mathematical tasks (design), and have student engagement with reasoning (content engagement). Based on Figure 1, this model necessitates a strongly committed teacher, but it also presents a reasonable responsibility for students to engage with each other in productive discourse and engage with the content. The strength of such a classroom model embeds discourse as a major component for both student and teacher to be respectful, centered on student learning, and assist with creating equitable spaces for learning opportunities (Boaler, 2006; Hiebert & Grouws, 2007; Yackel & Cobb, 1996).

Our research team recognizes existing work in the development of researched-based observation protocols. The IQA, MQI, M-SCAN, and RTOP serve specific purposes for research and practice. Each has specific strengths in what they focus on measuring, but they each possess some weaknesses as identified in the scholarly paper of Boston, Bostic, Lesseig, & Sherman (2015).The development of the MCOP^2 instrument design was to specifically fill gaps and improve on areas of weaknesses of these existing protocols in the field. More specifically, we designed the instrument to be "math teacher accessible" with little to no training needed but expectations of practice using the protocol in a low stakes environment among collaborative peers. Moreover, the MCOP^2 presents an opportunity for teachers and their peers to utilize a reasonable amount of time to work in partnership to improve the teaching and learning of mathematics based on what more than 125 experts in mathematics education externally validated as the important aspects of the mathematics classroom for teaching for mathematical understanding. The MCOP^2's major strength is the fact that these field experts validated the content of the instrument, and then 26 of these experts re-validated the revised and final version. The external validation gives strength with major agreement across the field.

THE MCOP² INSTRUMENT STATISTICAL ANALYSES

Utilizing a balanced data collection of classroom observations in elementary, secondary, and tertiary classrooms, the instrument revealed two measurable factors of teacher facilitation (α=0.850) and student engagement (α=0.897) consisting of seven items the load onto a single factor, as well as two items that load onto both factors. The two factors were confirmed via a Scree Plot (See Figure 2) and an Exploratory Factor Analysis (EFA). Six individuals with varying backgrounds in mathematics education and without discussion scored classroom videos from K-2, 3-5, 6-8, 9-12, and tertiary mathematics classrooms. The results indicated intra-class correlations (ICC) of 0.669 on the student engagement scale and 0.616 on the teacher facilitation score. Based on the fact that these six individuals could achieve an ICC without discussion and with isolated observational scoring, presents an acceptable ICC and high degree of agreement on each factor (Cicchetti, 1994; Lawshe, 1975). With additional practice and discussion with the instrument, we have seen the ICCs among users increase significantly.

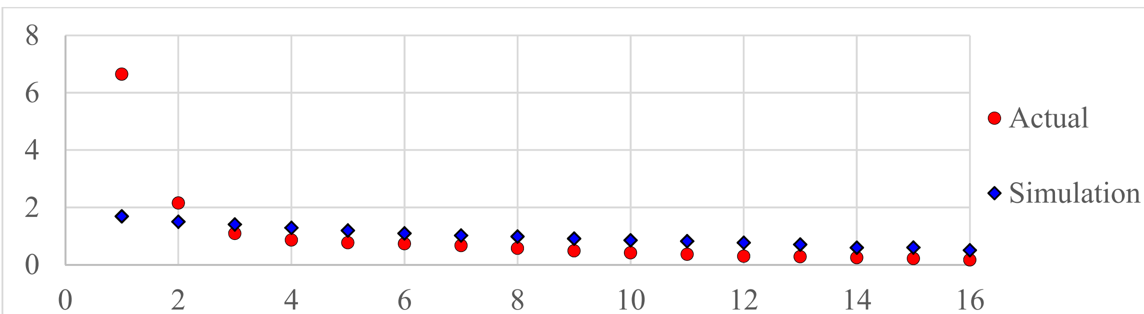

Figure 2. Scree Plot of MCOP² Internal Structure Analysis

Figure 3 presents the 16 items of the MCOP² and their respective loadings on each factor. Each item scores on a 0, 1, 2, or 3 point scale for which a specific rubric description indicates behaviors and characteristics in the classroom for the teacher and student to score at each point level.

Teacher Facilitation for Teaching Mathematics		Student Engagement in Learning Mathematics	
.851	The lesson involved fundamental concepts of the subject to promote relational/conceptual understanding.	Students engaged in exploration/ investigation/ problem solving.	.783
.503	The lesson promoted modeling with mathematics.	Students used a variety of means (models, drawings, graphs, concrete materials, manipulatives, etc.) to represent concepts.	.612
.585	The lesson provided opportunities to examine mathematical structure. (symbolic notation, patterns, generalizations, conjectures, etc.)	Students were engaged in mathematical activities.	.873

.508	The lesson included task(s) that have multiple paths to a solution or multiple solutions.	There were a high proportion of students talking related to mathematics.	.831
.568	The lesson promoted precision of mathematical language.	Students persevered in problem solving.	.737
.744	The teacher's talk encouraged student thinking.	In general, the teacher provided wait-time.	.422
.738	The teacher uses student questions/comments to enhance mathematical understanding.	Students were involved in the communication of their ideas to others (peer-to-peer).	.863
.410	Students critically assessed mathematical strategies.		.359
.329	There was a climate of respect for what others had to say.		.397

Figure 3. MCOP^2 Items and Respective Factor Loadings

Example rubrics for one item on each scale

One example from the student engagement factor we have selected to present at ICME for discussion is the *students used a variety of means to represent concepts* seen in Figure 4. This item selection presents one example of how the focus of the observation is on what students do with the content to improve their conceptual understanding.

3	The students manipulated or generated two or more representations to represent the same concept, and the connections across the various representations, relationships of the representations to the underlying concept, and applicability or the efficiency of the representations were explicitly discussed by the teacher or students, as appropriate.
2	The students manipulated or generated two or more representations to represent the same concept, but the connections across the various representations, relationships of the representations to the underlying concept, and applicability or the efficiency of the representations were not explicitly discussed by the teacher or students.
1	The students manipulated or generated one representation of a concept.
0	There were either no representations included in the lesson, or representations were included but were exclusively manipulated and used by the teacher. If the students only watched the teacher manipulate the representation and did not interact with a representation themselves, it should be scored a 0.

Figure 4. Student Engagement Rubric Item

Additionally, we selected the teacher facilitation item for discussion as the *The Teacher's Talk Encouraged Student Thinking* (see Figure 5) as it relates well with respect to focusing on teaching for conceptual understanding.

3	The teacher's talk focused on **high levels** of mathematical thinking. The teacher may ask lower level questions within the lesson, but this is not the focus of the practice. There are three possibilities for high levels of thinking: analysis, synthesis, and evaluation. **Analysis**: examines/ interprets the pattern, order or relationship of the mathematics; parts of the form of thinking. **Synthesis**: requires original, creative thinking. **Evaluation**: makes a judgment of good or bad, right or wrong, according to the standards he/she values.
2	The teacher's talk focused on **mid-levels** of mathematical thinking. **Interpretation**: discovers relationships among facts, generalizations, definitions, values and skills. **Application**: requires identification and selection and use of appropriate generalizations and skills
1	Teacher talk consists of "**lower order**" knowledge based questions and responses focusing on recall of facts. **Memory**: recalls or memorizes information. **Translation**: changes information into a different symbolic form or situation.

0	Any questions/ responses of the teacher related to mathematical ideas were rhetorical in that there was no expectation of a response from the students.

Figure 5. Teacher Facilitation Rubric Item

DISCUSSION

Theoretically, an observation may score very well on the student engagement factor with a lesson not focused on conceptual understanding (e.g. the teacher presents examples focused on procedural skill, followed by the students imitating that skill on several problems working collaboratively in groups). On the other hand, a lesson focused heavily on conceptual understanding may score high on the teacher facilitation factor, but students may forsake to engage in their accountabilities (e.g. a teacher gives an excellent presentation of mathematical modeling, focusing on structure, with little expectation of students engaging in problem solving or discussion). Such results may correspond to a classroom not suitable within the Community of Learners framework.

Lessons built around the paradigm of direct instruction, as defined in Munter, Stein, and Smith (2015), would likely result in low scores on two items, due to focus on student talk that is not a major focus in the direct instruction paradigm. Another item focusing on multiple paths to a solution or multiple solutions does not fit well within the direct instruction paradigm as the paradigm does not have as much emphasis on creativity. The remaining items, however, do fit well within the direct instruction paradigm. Such lessons would likely have slightly lower scores on the Student Engagement subscale with slightly higher scores on the Teacher Facilitation subscale.

A lesson implemented within the framework of dialogic teaching (Munter, Stein, & Smith, 2015) would result in a higher emphasis on student authority and responsibility. Lessons implemented within this dialogic teaching paradigm would likely receive a high score on the Student Engagement subscale. This could lead to a slightly lower score on the Teacher Facilitation subscale if the teacher had much less of a role during a well-engaged set of students in the lesson. Therefore, it is likely that the scores on the $MCOP^2$ subscales may be influenced by the relationship of the teaching paradigm in the classroom between direct instruction and dialogic teaching.

When choosing an instrument for research involving observations of mathematics teaching, the $MCOP^2$ stands apart in that it measures the actions of both the teacher and students. By capturing both sets of actions, the $MCOP^2$ differentiates between teacher design and implementation of a lesson and the students' engagement with peers and teacher. This makes the $MCOP^2$ a holistic assessment of the classroom environment not found in other observation protocols. To accompany the instrument, a user guide was developed to clarify and well-define the 16 indicators while providing explicit details regarding the levels of scoring. This eliminates the need for extensive training with the assumption that the user understands the terminology and concepts in the scoring rubrics. It is recommended that the observer stay for the entirety of the lesson to account for teacher design and structure of the lesson, and student interactions.

The $MCOP^2$ was not designed to evaluate a teacher on a single observation due to the nature and complexity of the teaching, but to be used for multiple teaching episodes (three to six we recommend) to adequately capture the instructional practice annually. We understand the instructional decisions necessary to effectively teach students differ on a daily basis. This includes a teacher who realizes that a more direct instruction approach might be necessary following days of student centered investigations. Or if the teacher is working on procedural fluency, those lessons may not score as high and would not be representative of the teacher's typical instructional practice.

We recommend using the MCOP^2 in a pre-post design as part of professional development at a very minimum with the possibility of observations during ongoing professional development. This allows teachers to see their baseline highlighting the practices that are strong within their teaching practice and those that are not as apparent, and then through professional development the teachers can work to strengthen those practices. Our research team envisions teachers within schools and/or districts using the MCOP^2 as a means of peer observations to learn about themselves and their colleagues to improve site-wide practice to improve the teaching and learning of mathematics.

At ICME13, we present the MCOP^2 to the international community for use in translation research to additional languages and cultures, as well as the specificity of the rubric descriptors of the 16-item protocol. Moreover, the use of the MCOP^2 has been widely accepted as part of the United States Mathematics Teacher Education Partnership (MTE-P) that is part of the Association of Public and Land-grant Universities (APLU) Science and Mathematics Teacher Imperative (SMTI). The partnership has more than 100 universities and colleges with many using the instrument for data collection in their teacher preparation programs for a national sample to help facilitate the national partnership's ability to improve programs and the quality of beginning teachers of mathematics with these essential practices of teaching and learning mathematics.

Our team's instruction book, the MCOP^2 long and short forms, can be found at the following URL: http://jgleason.people.ua.edu/mcop2.html

References

Ball, D.L., Ferrini-Mundy, J., Kilpatrick, J., Milgram, R.J., Schmid, W., & Schaar, R. (2005). Reaching for common ground in K-12 mathematics education. Notices of the American Mathematical Society, 52(9), 1055-1058.

Barker, W., Bressoud, D., Epp, S., Ganter, S., Haver, B., & Pollatsek, H. (2004). Undergraduate programs and courses in the mathematical sciences: CUPM curriculum guide. Washington, DC: Mathematical Association of America.

Boaler, J. (2006). How a detracked mathematics approach promoted respect, responsibility, and high achievement. Theory into Practice, 45(1), 40-46.

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

Brownell, W. A. (1935). Psychological considerations in the learning and teaching of arithmetic. In W. D. Reeve (Ed.), The teaching of arithmetic: Tenth yearbook of the National Council of Teachers of Mathematics (pp. 1–31). New York: Teachers College, Columbia University.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6(4), 284-290.

Cobb, P., Wood, T., & Yackel, E. (1993). Discourse, mathematical thinking, and classroom practice. In E.A. Forman, N. Minick, & C.A. Stone (Eds.), C*ontexts for learning: Sociocultural dynamics in children's development* (pp.91-119). Oxford: Oxford University Press.

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

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 65–97). New York: Macmillan.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), Second handbook of research on mathematics teaching and learning (pp. 371-404). Reston, VA: National Council of Teachers of Mathematics.

Lawshe, C. H. (1975). A quantitative approach to content validity. Personnel Psychology, 28, 563-575.

Munter, C., Stein, M.K., & Smith, M.S. (2015). Dialogic and direct instruction: Two distinct models of mathematics instruction and the debate(s) surrounding them. Teachers College Record. To appear.

Nathan, M. J., & Knuth, E. J. (2003). A study of whole classroom mathematical discourse and teacher change. Cognition and instruction, 21(2), 175-207.

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

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

National Council of Teachers of Mathematics. (2006). Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence. Reston, VA: Author.

National Council of Teachers of Mathematics. (2011). Focus in high school mathematics: Fostering reasoning and sense making for all students. M. E. Strutchens & J. R. Quander (Eds.). Reston, VA, Author.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common core state standards mathematics.* Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers. Retrieved from http://www.corestandards.org/Math

Polya, G. (1957). Mathematical discovery: on understanding, learning, and teaching problem solving. (2nd edition). Princeton, NJ: Princeton University Press.

Rogoff, B., Matusov, E., & White, C. (1996). Models of teaching and learning: Participation in a community of learners. In D. R. Olson & N. Torrance (Eds.), The handbook of education and human development (pp. 388–414). Cambridge, MA: Blackwell Publishers.

Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2010). Measuring reform practices in science and mathematics classrooms: The reformed teaching observation protocol. *School Science and Mathematics, 102*(6), 245-253.

Skemp, R. R. (1971). The psychology of learning mathematics. Middlesex, England: Penguin.

Skemp, R. R. (1976). Relational understanding and instrumental understanding. Arithmetic Teacher, 26(3), 9–15.

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

Stein, M.K., Smith, M.S., Henningsen, M.A., & Silver, E.A. (2009). Implementing standards-based mathematics instruction: A casebook for professional development (2nd edition). New York, NY: Teachers College Press.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. Journal for Research in Mathematics Education, 27(4), 458-477.