Using the MCOP^2 as a Grade Bearing Assessment of Clinical Field Observations

Jeremy Zelkowski
The University of Alabama
jzelkowski@ua.edu

Jim Gleason
The University of Alabama
jgleason@ua.edu

The development and validation of observation protocols for lessons implemented in mathematics classrooms is scarce. That is to say, protocols for mentor teachers or university supervisors to use when observing and scoring preservice mathematics teachers are virtually absent in the field of mathematics teacher preparation. Departmental or program checklists, aligned to discipline-agnostic state teaching standards, are more commonly used for preservice teacher observations (Brown & Crippen, 2016; Caughlan & Jiang, 2014; Morrell, Flick, & Wainwright, 2004). Because many secondary mathematics teacher preparation programs (SEMA-TPP) are housed in departments of education, secondary mathematics educators located in a Mathematics department may have little flexibility with regards to the documentation of generalist practices, dispositions, and standards common across disciplines. Further, the accreditation *unit* is often the Secondary Teacher Preparation Program rather than the more discipline-specific SEMA-TPPs. To add to this complication, most SEMA-TPP's are small, providing limited opportunity to collect large-scale data in a reasonable amount of time to do rigorous quantitative analyses.

Within a similar context, the University of Alabama (UA) mathematics education faculty developed a research-based observation protocol for K-16 mathematics classrooms usable for cooperating mathematics mentor teachers, supervisors, and faculty. The Mathematics Classroom Observation Protocol for Practices (MCOP^2) was developed during 2012-15 to evaluate a large, grant-funded professional development project. The research on the development of MCOP^2, validation, reliability, and factor analyses resulted in strong findings. Without training, good interrater reliability on the two factors was determined (IRR=0.669 on student engagement [SE], IRR=0.616 on teacher facilitation [TF]). On the two factors, SE had an α=0.897 and the TF had an α=0.850. Finally, the two national surveys of experts in the field produced a strong external validation of the MCOP^2 (Gleason, Livers, & Zelkowski, 2017).

This brief research paper presents the results from a two-year study in the UA SEMA-TPP to align grading practices of the MCOP^2 with previous un-validated generalist observation forms. The generalist observation forms have historically been used at UA to address state teacher preparation standards, as well as to determine letter grades assigned to lessons

Lawler, B. R., Ronau, R. N., & Mohr-Schroeder, M. J. (Eds.). (2016). *Proceedings of the fifth annual Mathematics Teacher Education Partnership conference*. Washington, DC: Association of Public Land-grant Universities.

conducted by grades 6-12 preservice mathematics teachers. We set out to determine: (1) What grading scale would be appropriate on the MCOP^2 and correlate well to the grades received on the generalist observation form? (2) How do preservice teachers perceive the MCOP^2 as their grade-determining observation protocol, as opposed to the generalist observation form? (3) Does preparing preservice teachers to consider the MCOP^2 in their planning of formal observation lessons improve the quality of the lessons than prior to the MCOP^2 implementation? (4) What are the impacts for both students and program to use the generalist form as a non-grade-bearing, end-of-semester, summative form, and the MCOP^2 as the exclusive grade-determining observation protocol for both methods students and student teachers?

Institutional Transformation of SEMA-TPP to the College

In 2008, the first author arrived at UA to lead the SEMA-TPP and quickly joined forces with the second author, a mathematician responsible for the curriculum of the secondary mathematics major track. A partnership was created and the secondary mathematics education major track was aligned to the recommendations of *The Mathematical Education of Teachers* (Conference Board of the Mathematical Sciences [CBMS], 2001), and later *The Mathematical Education of Teachers II* (CBMS, 2012) ensued. Changes included developing a second capstone course, revising the first capstone course, inserting analysis and differential equations as requirements, and adjusting the programming ancillary course. On the education side, three methods courses were aligned sequentially alongside the capstone math courses, and structured to create a two-year cohort model to develop habits of mind and mathematical practices that relate to teaching mathematics with quality tasks in a manner that maintains high level cognitive demand (Stein, Smith, Henningsen, & Silver, 2009).

A few years later, the development and use of the MCOP^2 spurred the entire college to begin transforming other secondary education programs, with Social Studies and English programs utilizing structures similar to mathematics. These initial transformations led to development of a third capstone course that has since been piloted twice and will become permanent in the Fall of 2017. The transformation of the SEMA-TPP since 2008, and more recently with the MCOP^2, has moved the preparation program to a point that the *quality* of our graduates and their ability to model best practices *regularly* has out-performed an old philosophy in which the *quantity* of graduates was of greater concern. Initially, resistance came administratively, but the program rigor, voluntary NCTM SPA accreditation review, and research-based design convinced other disciplines to transform. Both program quality and quantity have increased without reduction of rigor.

Theoretical Framework for the MCOP²

The MCOP² became an important mechanism to the SEMA-TPP to communicate expectations for mathematics instruction, to provide feedback to secondary mathematics students as they designed and implemented lessons, and to evaluate the results of the program. The MCOP2 was designed to reflect the important qualities of mathematics instruction as advised by the profession. For example, conceptual understanding focuses on students drawing, inferring, and ultimately making connections between mathematical operations, representations, structure, reasoning, and modeling through engagement with content and through discourse (Brownell, 1935; Hiebert & Carpenter, 1992; Hiebert & Grouws, 2007; National Research Council, 2003 Resnick & Ford, 1981; Skemp, 1971, 1976). With that said, a preservice teacher's lesson would not fit well within a goal for conceptual understanding if the lesson focused solely procedural fluency and efficiency through instructional practices such as drill worksheets emphasizing procedures to be learned by rote. This practice may be beneficial at times within the larger context of mathematics instruction, though should be neither the primary method of instruction; it fails to engage students in doing and understanding mathematics beyond procedural fluency. Furthermore, "there is no reason to believe, based on empirical findings or theoretical arguments, that a single method of teaching is the most effective for achieving all types of learning goals" (Hiebert & Grouws; 2007, p. 374).

From the beginning, Cohen, Raudenbush, & Ball's (2003) theoretical model of instruction as interaction between teachers, students, and mathematical content guided the development of the MCOP² framework of teaching for conceptual understanding through various forms of instructional practices. The team began with examining the commonly used Reformed Teaching Observation Protocol (Sawada et al., 2010) as well as the Standards for Mathematical Practice (SMPs) from the Common Core State Standards, which have been shown to align well with the NCTM process standards (Koestler, Felton, Bieda, & Otten, 2013). We then considered the UA SEMA-TPP program standards, well-aligned to the NCTM Specialized Professional Association (SPA) standards (NCTM, 2012), and assigned connections between these standards and the Cohen et al. (2003) model. As seen in Figure 1, the Instruction as Interaction model developed at UA presents three nodes, or factors, which could be denoted as teacher responsibilities, mathematical content of the lesson, and student engagement.

At the completion of the aforementioned factor analysis (Gleason et al., 2017), a two-factor structure of the MCOP² was found, as opposed to the initial design of three factors seen in the Instruction as Interaction model. The final instrument more closely aligns with Rogoff, Matusov, and White (1996), a framework characterized by a community of learners in which authority and responsibility are shared among students and teachers. This is evident in our model (Figure 1) by visualizing both *halves* of the oval as denoted with a dotted line. The teacher provides goals and objectives of the lesson while targeting learning and problem solving through discourse. Likewise, the role of students to engage in the mathematical content of the lesson through problem solving and discourse is a shared responsibility to nurture both sides of responsibility that will increase the likelihood of mathematical learning. The MCOP^2 framework and measurable factors emphasize that both teacher and student have shared responsibilities within the learning community of the mathematics classroom.

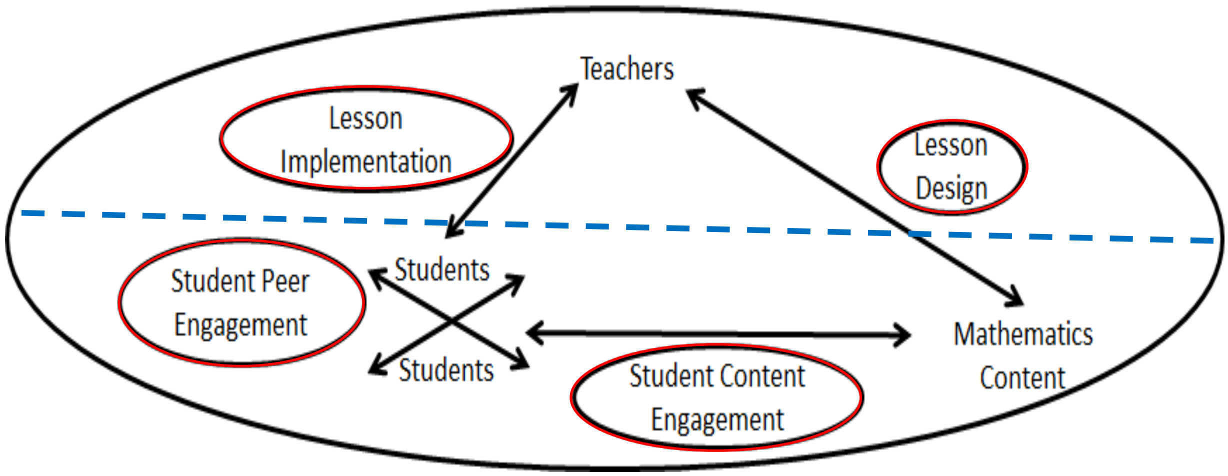

Figure 1. Instruction as Interaction model (Based upon Cohen et al., 2003).

Methodology

The SEMA-TPP aimed to replace a generalist secondary programs form for grade-bearing observations of methods students and teacher with a specific mathematics observation protocol targeting both teacher facilitation and student engagement measures of effective instruction.[1] To support this shift and subsequently address the research questions, the SEMA-TPP collected observational data and survey data for two years, during the 2014-15 and 2015-16 academic calendars. Observations of the delivery of mathematics lessons in 6-12 mathematics classrooms by methods students and student teachers used both observation forms, with the order of form completion alternating between observations. This method was used to increase the reliability between scores on both forms, as well as increase the validity of accurately recording the observation across both forms (Bryman, 2012). The generalist form was used in the determination of student grades, and although MCOP^2 data was collected, it did not impact grades. At the conclusion of Year 1, data from the paired observation forms was presented to the secondary programs faculty. SEMA-TPP faculty moved forward in using each protocol in Year 2 with each form counting as 50% of the observational grades.

[1] The generalist form in use at UA can be found at http://bit.ly/SecObsForm, and the MCOP^2 can be found at http://bit.ly/UA-MCOPP.

Brief Overview of the Results

We address each research question individually. The first question considered (1) What grading scale would be appropriate on the MCOP^2 and correlate well to the grades received on the generalist observation form? Over the two-year study, we examined 59 SEMA-TPP candidates in middle and upper grades mathematics classrooms using both observation forms. Of the 59 candidates, 31 were students enrolled in a methods course, while 28 were student teachers in the subsequent and final term in the program. Analysis of the data yielded a *very strong* correlation between scores on the two forms, shown in Figure 2.

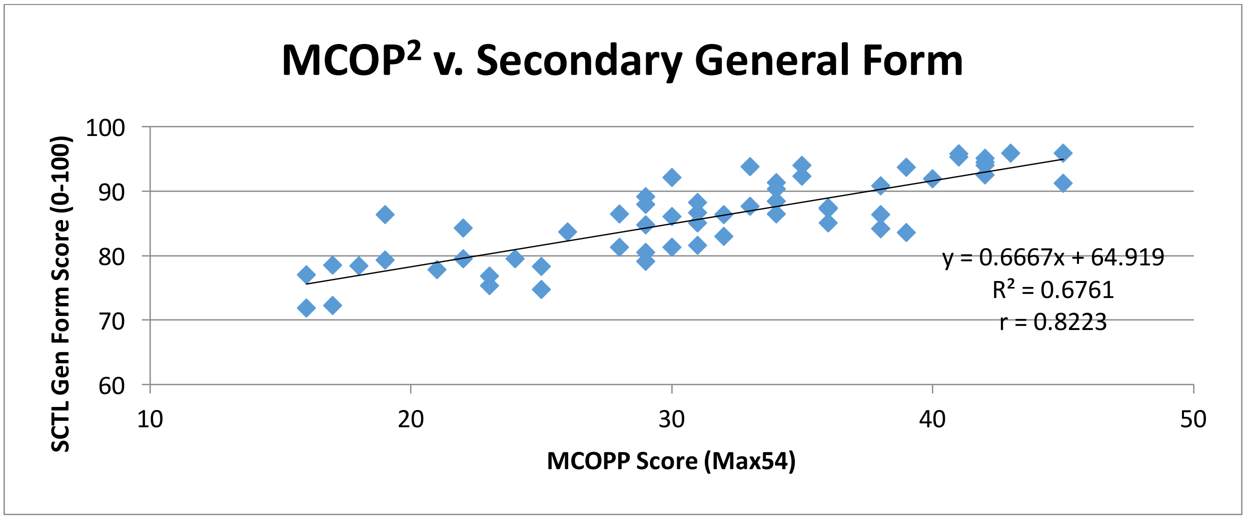

Figure 2. Correlational analysis of the MCOP^2 observation form to the generalist form.

Such a result is sufficient evidence to warrant using either form. On the other hand, this result is heavily reliant on one rater; the rater in this case is the lead author who regularly scored lessons on the generalist form in the C or average range (range during study 72 to 97). Cooperating mathematics teachers, in contrast, rarely scored observations of student teachers less than a mid-B, mostly assigning A's (95% of the time). Thus, the strong correlation of the findings indicate that an accurately scored MCOP^2 rubric will align to A, B, C, D, F letter grades easily, whether used by a teacher, supervisor, or university faculty. Next was to consider what MCOP^2 point total equates to each letter grade A, B, C, D, and F.

In Figure 3, we examine the use of the linear model in Figure 2, while in Figure 4 we present an improved spread across excellence, good, average, and below average (failing) scores. Using only the linear regression model on the MCOP^2 point totals, Figure 3 reveals grades would convert to a range of about 75 to 95. This result aligns to the range of scores that supervising teachers regularly assigned, 80-100, on the generalist forms reduced by five points. Therefore, we decided to examine the MCOP^2 means per item (0.0 to 3.0).

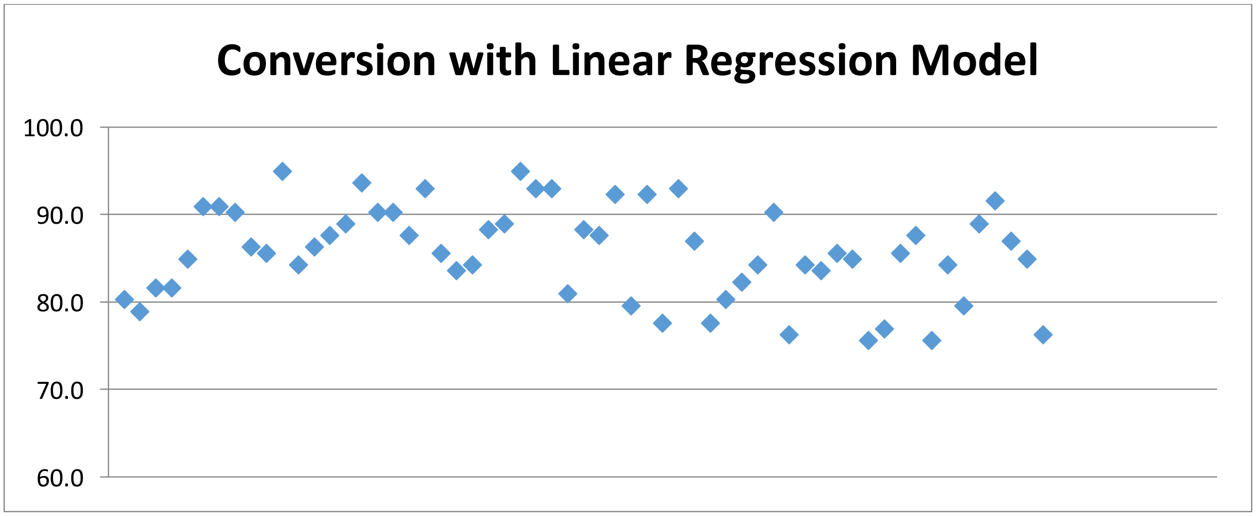

Figure 3. Conversion of MCOP^2 scores using least squares linear model.

We explored a few options and settled on a conversion that is shown in Figure 4 to obtain the desired distribution. Figure 4 demonstrates that the spread of the converted scores, 68 to 100, is very similar in nature to the spread of generalist form scores by the lead author, 72 to 97.

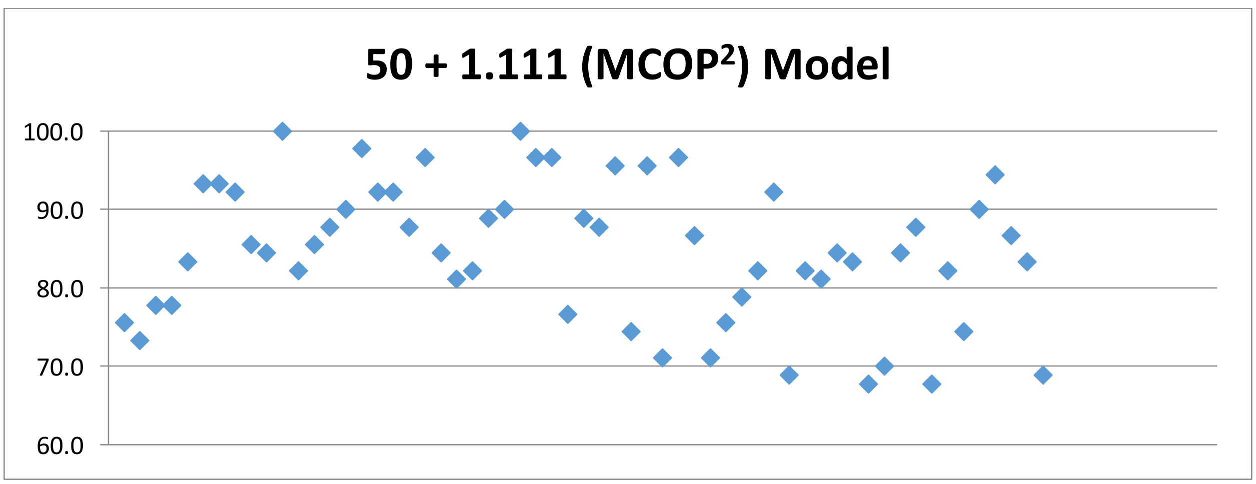

Figure 4. UA SEMA-TPP adopted grade conversion model for the MCOP^2. Notes on conversion: MCOP^2 mean score of 2.5 is equivalent to 100%, an A+; MCOP^2 mean score of 1.0 is equivalent to 70%, a C-.

The raw scores from the MCOP^2 can range in total points from zero to 54. It is unreasonable to expect perfect threes, let alone half twos and half threes from novice teachers. However, programmatically we view a mean of 2.5 as a top score for a preservice teacher designed and delivered lesson and a mean of 1.0 as minimally passing. The range 1.0 to 2.5 on the MCOP^2 maps to a distribution from the mid-60s (failing) to 100 (perfect) for converting to a grade. We have adopted the linear score conversion of *grade*=50+10/9**MCOP^2* (shown in Figure 4) to translate the MCOP2 score to a grade out of 100 points, revealing excellence, good, average, and in need of improvement to pass.

Our second research question was, (2) How do preservice teachers perceive the MCOP^2 as their grade-determining observation protocol, as opposed to the generalist observation form? The resulting MCOP^2 grades were higher or almost equal than the general form grades about 2/3 of the time. Eighty-three percent (n=49 of 59) of preservice teachers reported, through online anonymous surveys, they preferred the MCOP^2 score because they knew what to improve on for the next observation whereas the generalist form was not specific enough without written feedback or discussion. This feedback was very telling; they preferred knowing exactly their role and their students' roles in a formally observed lesson as it was outlined in the MCOP^2. A large portion (n=25 of 28) said they preferred the MCOP^2 during their student teaching so that *only mathematics teachers* would score and determine their grades. About half of student teachers are observed by out-of-field teachers within schools for upwards of half their observational scores during their student teaching internship. This practice has since changed to only formative non-grade bearing feedback by out-of-field teachers. MTE-P partners have indicated this out-of-field observation practice is common even though it does not provide feedback to the preservice teachers or program faculty regarding the implementation of high quality mathematics lessons.

We also set out to study, (3) Does preparing preservice teachers to consider the MCOP^2 in their planning of formal observation lessons improve the quality of the lessons than prior to the MCOP^2 implementation? To answer this, we examined the *lesson plan* scores of all formal observations during the two years of data collection of all methods students (student teachers were not scored on lesson plans). We then looked back to the prior two years of lesson plan grades of generalist form scored observations of methods students of lessons implemented. We found that 26 of 31 (84%) method student lesson plans from two years of MCOP^2 were scored higher than in the two years prior to MCOP^2. This indicates a higher quality in lesson planning when the MCOP^2 is used as an observation protocol. When comparing the pre-MCOP^2 era observation scores on the generalist form to the observations with the generalist form during the MCOP^2 use, the same 84% of observations scored better. Hence, we conclude that both planning and enactment was clearly better with the MCOP^2 use in our program.

Conclusion

Our conclusions of this preliminary work to begin and understand the implementation of the MCOP^2 in the UA SEMA-TPP program reside in our fourth research question, (4) What are the impacts for both students and program to use the generalist form as a non-grade-bearing, end-of-semester, summative form, and the MCOP^2 as the exclusive grade-determining observation protocol for both methods students and student teachers? A unit, college, or program can and should use generalist summative tools for analyzing large-scale factors as overall university/college preparation, secondary programs preparation, or state-level preparation. On the other hand, at the level in which we evaluate and assess our teacher candidates during methods and student teaching, the $MCOP^2$ results clearly promote better designed lessons, and stronger implementation of those lessons. In addition, teacher candidates prefer the $MCOP^2$ to help because of the detailed information provided to guide their professional development and the more accurate determine their grades for formal observations of teaching.

Our institutional colleagues have recognized this work and are now following on the heels of the successes of the SEMA-TPP. More impacted are the local cooperating teachers who now find great value in a well-defined rubric measuring both sides of their responsibilities, the mathematics classroom itself, and supporting the student teacher to achieve this target. Moreover, the cooperating teachers do not feel pressured to give higher grades. Some teachers have mentioned using the $MCOP^2$ helps them now look for certain practices of themselves and their students in their normal classroom daily operation. This impact on cooperating teacher instructional practices implicitly from classroom teachers using the $MCOP^2$ to evaluate preservice teachers is an intriguing area for future research.

Drawbacks and Limitations

We recognize not all programs can implement the practices at UA as many out-of-field teachers evaluate methods students or student teachers, and some programs do not have a year of strong preparatory content, pedagogy, and clinical field experiences before methods and student teaching. These practices for preparation of secondary mathematics teachers at UA are our implementation of the recommendations of MET and MET II, structures intended to best serve both preservice teachers and the students in secondary schools in which they will teach. We are concerned that the practice of evaluation done by out-of-field educators works against the efforts to transform mathematics teacher preparation in the eyes of the public and stakeholders. High stakes testing like the edTPA are here because the field of teacher preparation has allowed a perception to emerge that we produce an unacceptable percentage of unprepared and unqualified credentialed teachers. We question the ethical rationale of any practice that contributes to the public perception of weak teacher preparation, a problem heightened in mathematics due to the generally negative views of mathematics in the U.S.

Next Steps

Our next steps on our program transformation at UA SEMA-TPP center around three items. First, we aim to begin work on tying student achievement measures to the $MCOP^2$. This is a massive endeavor requiring funding for a large-scale study. Second, we aim to begin analyzing our NCTM SPA data set to extract programmatic variables that influence $MCOP^2$ scores and make changes, additions, and deletions to the programmatic structure where needed. Lastly, we aim to begin work with our cooperating and mentor teachers with the $MCOP^2$ as previously mentioned.

We hope that the MTE-P community welcomes this instrument and finds it useful in aiding the similar transformation of their secondary mathematics teacher preparation program, especially for evaluating methods students and student teachers in the mathematics classroom.

References

Brown, J. C., & Crippen, K. J. (2016). The growing awareness inventory: Building capacity for culturally responsive science and mathematics with a structured observation protocol. *School Science & Mathematics*, *116*(3), 127-138.

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.

Bryman, A. (2012). *Social research methods (4th edition)*. Oxford University Press: Oxford, UK.

Caughlan, S., & Jiang, H. (2014). Observation and teacher quality: Critical analysis of observational instruments in preservice teacher performance assessment. *Journal of Teacher Education*, *65*(5), 375-388.

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

Conference Board of the Mathematical Sciences (2001). *The mathematical education of teachers.* Providence, RI and Washington, DC: American Mathematical Society and the Mathematical Association of America.

Conference Board of the Mathematical Sciences (2012). *The mathematical education of teachers II.* Providence, RI and Washington, DC: American Mathematical Society and the Mathematical Association of America.

Gleason, J.A., Livers, S.D., & Zelkowski, J. (2017). Mathematics Classroom Observation Protocol for Practices (MCOP^2): A validation study. *Investigations in Mathematics Learning, TBD.*

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.

Koestler, C., Felton, M.D., Bieda, K.D., & Otten, S. (2013). *Connecting the NCTM process standards and the CCSSM practices.* Reston, VA: NCTM.

Morrell, P. D., Flick, L., & Wainwright, C. (2004). Reform teaching strategies used by student teachers. *School Science & Mathematics*, *104*(5), 199-209.

National Research Council. (2003). Evaluating and improving undergraduate teaching in science, technology, engineering, and mathematics. Washington, DC: The National Academies Press.

National Council of Teachers of Mathematics (NCTM). (2012). *Specialized Professional Association (SPA) 2012 Standards for the Council for the Accreditation of Educator Preparation (CAEP).* Reston, VA: Author. Retrieved online from www.nctm.org/Standards-and-Positions/CAEP-Standards

Resnick, L. B., & Ford, W. W. (1981). *The psychology of mathematics for instruction.* Hillsdale, NJ: Erlbaum.

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, March 17). 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., 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.