# MATHEMATICS CLASSROOM OBSERVATION PROTOCOL FOR PRACTICES RESULTS IN UNDERGRADUATE MATHEMATICS CLASSROOMS

Jim Gleason
The University of Alabama

Leah D. Cofer
The University of Alabama

*The purpose of this study is to determine the reliability and validity of the Mathematics Classroom Observation Protocol for Practices (MCOP$^2$) in undergraduate mathematics classrooms, an observation instrument designed to measure the degree to which a mathematics classroom aligns with the standards put forth by national mathematics organizations. To examine the reliability and validity of the MCOP$^2$ in the undergraduate setting, over thirty undergraduate mathematics classrooms at a large southeastern university were observed during the fall semester of 2013. The exploratory factor analysis conducted from the data collected indicates there are two main factors to consider in an undergraduate mathematics classroom: "lesson content" and "student engagement and classroom discourse". The internal reliability of each of these factors was verified using classical test theory to measure well at the group level.*

*Key words:* Classroom Teaching, Evaluation, Standards

The Mathematics Classroom Observation Protocol for Practices (MCOP$^2$) is a K-16 mathematics classroom instrument designed to measure the degree of alignment of the mathematics classroom with the Standards for Mathematical Practice from the Common Core State Standards in Mathematics (NGACBP & CCSSO, 2010); "Crossroads" and "Beyond Crossroads" from the American Mathematical Association of Two-Year Colleges (AMATYC 1995; AMATYC 2006); the Committee on the Undergraduate Program in Mathematics Curriculum Guide from the Mathematical Association of America (Barker et al., 2004); and the Process Standards of the National Council of Teachers of Mathematics (NCTM, 2000). The instrument contains 17 items intended to measure three primary constructs (student engagement, lesson content, and classroom discourse) as validated by a review of over 150 individuals self-identified as mathematics teacher educators from a mixture of mathematics departments and departments or colleges of education (Gleason, Zelkowski, Livers, Dantzler, & Khalilian, 2014). Each of the 17 items also contains a full description of the item with specific requirements for each rating level; see Appendix B for sample items descriptions.

## Purpose and Proposed Uses

Using peer reviews to evaluate faculty members' teaching effectiveness is a policy that is currently "gaining momentum" in higher education (Harris, Farrell, Bell, Devlin, & James, 2008). Seldin (1999) best defends this type of evaluation in higher education by claiming a teacher's performance in the classroom should be considered comparable to their publications and thus held to the same review process. If a faculty member's peers are supposed to reliably evaluate his or her teaching performance through classroom observations, peer review requires the "essential ingredient" of a rating scale "with scale items (that) typically address the instructor's content knowledge, delivery, teaching methods, learning activities, and the like" (Berk, Naumann, & Appling, 2004). If observing faculty members use such a reliable classroom observational protocol, peer review has potential to better measure an instructor's teaching abilities than student evaluations since there are features of a lesson that peers are better qualified to evaluate than students (Harris et al., 2008; Berk, Naumann, & Appling, 2004).

There are many universities already utilizing some version of a classroom observation protocol during their faculty peer reviews. However, these preexisting protocols are very generic, lengthy, and subjective. For instance, University of New Mexico, Tallahassee Community College, and California State University, East Bay all have observation forms online that have been adapted from *A Guide for Evaluating Teaching for Promotion and Tenure* by Centra, Froh, Gray,& Lambert (1976) where the observing faculty member can fill out an extensive forty-five item form with only three potentially biased scoring options available: "not observed", "more emphasis", or "accomplished very well" (University of New Mexico, 2006; Tallahassee Community College, 2012; California State University, East Bay, 2013; Centra et al., 1976). Since these preexisting college level protocols are not subject-matter specific, they do not necessarily draw an observer's attention toward more specific aspects of the lesson, classroom, or students "thereby resulting in potentially different kinds of teacher evaluation practice" (Spillane, Halverson, & Diamond, 2001). Spillane, Halverson, & Diamond (2001) compare 'Protocol A' consisting of a checklist of generic teaching processes with a content-specific 'Protocol B' which includes items such as "how students were required to justify their mathematical ideas" to justify why a subject-matter specific instrument would allow faculty members to identify more precise details of a teacher's performance. Instead of a generic form, a reliable content-specific observational protocol like the $MCOP^2$ should be used during peer observations to measure a teacher's effectiveness and thus help generate a discussion on quality teaching in college undergraduate classrooms throughout the United States.

This review process is also useful for generating discussion among future and current college mathematics faculty about teaching. It can be used to help new graduate students better prepare for the classroom, help departments to decide goals for teaching and how well those goals are being met, and be used while observing classes as a group to generate discussion about what the standards look like in the college setting.

Since the $MCOP^2$ is grounded in the national recommendations of organizations focused on post-secondary education, it is useful to explore the current practice of mathematics teaching and its relationship to student learning. This instrument allows for a quantification of different aspects of college mathematics teaching that could then be used to explore teachers' choices in the classroom, effects of different teaching styles with different types of students, how a teacher's practice in the classroom changes with the different topics and situations throughout a semester, and many more.

## Preexisting Protocols

There are many content-specific classroom observation protocols already available for use in elementary, middle, and high school mathematics classes (Hill, Charalambous, Blazar, et al., 2012). Three existing classroom protocols claim to extend to college level mathematics classrooms, and while each of these protocols is described as unique, all three credit Horizon Research Corporation, Inc. for their development (Weiss, Pasley, Smith, Banilower, & Heck, 2003; Wainwright, Morrell, Flick, & Schepige, 2004; Walkington et al., 2012; Sawada et al., 2000a). Unlike the $MCOP^2$, these preexisting protocols are not designed specifically for mathematics classrooms, but instead are intended for use in both mathematics and science classrooms (Wainwright, Flick, & Morrell, 2003; Walkington et al., 2012; Sawada et al., 2000a). In order to maintain this dual purpose, logically these protocols use science terminology within some of their protocol descriptors such as "Students made predictions, estimations, and/or hypotheses and devised means for testing them" (Sawada et al, 2000b), making it difficult for an observer of a mathematics classroom to definitively score certain items. While some of the preexisting protocols have claimed to test for predictive validity in college mathematics classrooms (Sawada et al., 2000b), no preexisting protocol has done a study including more than a few strictly undergraduate mathematics classrooms so no preexisting protocol has actually proven its reliability or validity in college mathematics classrooms.

One of the most widely used classroom observational protocols in public school mathematics classes is the Reformed Teaching Observation Protocol (RTOP), developed by the Evaluation Facilitation Group of the Arizona Collaborative for Excellence in the Preparation of Teachers (ACEPT) (Sawada et al., 2000a). While the RTOP is widely praised for its reliability and validity in both math and science public school classrooms, surprisingly the RTOP Reference Manual contains very few references to articles in mathematics education yet numerous references to articles in science education. Out of the seventeen references listed, six references are strictly for science education, four articles are on learning and the brain, two articles are the RTOP's first Technical Report and Training Guide, one article is from the 1980s on both math and science education, and the remaining four are citations to various years of NCTM standards from 1989 – 2000 (Sawada et al., 2000b). The creators of the RTOP admit in the "Test Development" section of their Reference Manual that the language of the items in the first draft of the instrument was "particularly referenced toward science teaching" and thus hard to interpret in a mathematics classroom (Sawada et al., 2000b). Mathematicians in the ACEPT project critiqued and suggested making "an unequivocal request to overhaul the science-dominated language"; therefore, a mathematics educator then modified the wording of items on the instrument without changing the original structure (Sawada et al., 2000b). While the RTOP in commonly used in both math and science public school classrooms, its references and item language make it better geared for use in science classes.

In addition, the instrument was originally tested in 13 "introductory" mathematics classes at universities and community colleges, but that is a substantially small portion of the 153 total classrooms that participated, particularly when compared to the 63 college level science classes observed (see Table 12, Sawada et al., 2000b). Furthermore, it is interesting to note that the college classroom teacher samples consisted of "a large number of faculty who were involved in the ACEPT initiative", thus the authors concluded this could be a reason why the college classrooms samples had higher scores on the RTOP than the middle and high school samples (Sawada et al., 2000b). In 2002, the ACEPT program tried to extend their method of reformed teaching to the college level by attempting to "incorporate reformed teaching methods in several nonmajors' and majors' courses"; hence, the RTOP was again tested in certain college classrooms. However, the only mathematics course observed was "Theory of Elementary Mathematics", a course designed specifically for preservice elementary school teachers (Lawson et al., 2002). Unlike the $MCOP^2$, the RTOP does not utilize the most recent national standards for mathematics classrooms and the RTOP has not extensively been tested in strictly mathematics college level classes taught by ordinary faculty members.

A classroom observation protocol that supposedly extends to college level mathematics classrooms is the Oregon-Teacher Observation Protocol (O-TOP), created by the Oregon Collaborative for Excellence in the Preparation of Teachers (OCEPT) as part of the Collaborative for Excellence in Teacher Preparation program of the National Science Foundation (Wainwright et al., 2003). According to Wainwright et al. (2004), "A major focus of the OCEPT grant was to engage science and mathematics faculty members teaching undergraduate courses in institutions across the state in a critical examination of their instructional practices." Even though the O-TOP was allegedly designed for use in both science and mathematics classrooms, their only citations strictly pertaining to mathematics are to various years of NCTM standards from 1989 – 2000 (Wainwright et al., 2004). Furthermore, the only mathematics classrooms observed were courses taught by OCEPT Faculty Fellows, so these were not typical college mathematics classes: "Of the 10 mathematics observations, two were lecture, one was lecture with discussion, and the remaining seven were small group discussion" (Wainwright et al., 2004). Despite its supposed reliability in Faculty Fellows mathematics classes, the OTOP's scientific nature and lack of recent mathematical standards make it undesirable for use in college mathematics courses.

Another classroom observation protocol supposedly appropriate for use in mathematics and science classrooms "from kindergarten to college" is the UTeach Observation Protocol (UTOP) created by the UTeach program at the University of Texas at Austin (Walkington et al., 2012). The protocol was developed to evaluate UTeach graduates, particularly Noyce Scholars, in order to fulfill a National Science Foundation requirement (Walkington & Marder, 2013). The language used within the UTOP demonstrates it is extremely science-based. For instance, an indicator in the protocol on the pace and flow of the lesson has a science-specific example listed with it in the Training Guide: "e.g. most of a science lab is focused on directions instead of content development" (Marder et al., 2010). Besides the science-specific language, another drawback to the UTOP is it is solely based off of NCTM standards from 1991 (Walkington et al., 2012). Furthermore, even though its authors originally planned to conduct observations at the college level in order to refine the instrument, no study has documented testing the UTOP in an undergraduate mathematics classroom (Walkington et al., 2012). The UTOP's lack of testing at the college level, along with its standards deficiency and overuse of scientific language, show its inapplicability to college mathematics classrooms.

## MCOP$^2$ Framework

Since the MCOP$^2$ was formed for mathematics-specific classrooms using Common Core State Standards in Mathematics (CCSSM) *Standards for Mathematical Practice*, the American Mathematical Association of Two-Year Colleges' *Crossroads* and *Beyond Crossroads*, the Mathematical Association of America's *CUPM Curriculum Guide*, and the latest NCTM standards, it is the only classroom observation protocol available that is applicable for use in K-16 mathematics instruction. After the nation adopted the CCSSM for use in public school mathematics classrooms, the Association of Public and Land-grant Universities (APLU) issued a brief laying out an "action agenda" with four main points for the role of higher education institutions (APLU, 2011). Point one addresses the issue of aligning curriculum between K-12 and higher education, and the APLU later indicates disciplinary departments should be "transforming introductory courses so that they are aligned with CCSSM (in both content and approach)" (APLU, 2011; King, 2011). College teachers themselves agree that their curriculum should be aligned with CCSSM since during a study on the applicability of the Common Core State Standards, over 1800 instructors found the Standards for Mathematical Practice to be extremely applicable and important to their courses (Conley, Drummond, de Gonzalez, Rooseboom, & Stout, 2011). Furthermore, point three of the APLU's agenda states higher education institutions should be "conducting research on issues of teaching and learning the Common Core State Standards, teacher quality, and the implementation of the Common Core State Standards" (APLU, 2011; King, 2011). Thus there is a need for a CCSSM-based mathematics classroom protocol and the MCOP$^2$ is the only protocol intentionally designed to meet this requirement.

Each of the items on the MCOP$^2$ was designed to coordinate with a Standard for Mathematical Practice, and in turn thus correlates to a recommendation in the CUPM Curriculum Guide. For instance, Item #9 on the protocol is "The lesson provided opportunities to examine elements of abstraction (symbolic notation, patterns, generalizations, conjectures, etc.)," matching the second Standard for Mathematical Practice that instructors should be aiming to teach their students**: "**CCSS.Math.Practice.MP2: Reason abstractly and quantitatively" (NGACBP & CCSSO, 2010). This concept also connects to Part 1 of the CUPM Curriculum Guide which gives recommendations for departments, programs, and all courses by Barker, et al. (2004): "For instance, one reason students encounter difficulty in applying mathematics to problems in other disciplines is that they have trouble identify appropriate mathematical procedures when problems are expressed with different symbols than those used in the mathematics classroom….instructors can go beyond conventional *x*, *y* notation to use a larger collection of symbols for both constants and variables." (p. 20)

Therefore, both the CCSSM and CUPM specifically address this important aspect of a teacher's lesson content which Item #9 is designed to measure. This correlation between the teacher and student behaviors detected by the MCOP$^2$, the Standards for Mathematical Practice, and the CUPM Curriculum Guide extends to all seventeen items on the protocol.

## Methodology

A pilot study to field test the MCOP$^2$ in undergraduate mathematics classrooms was implemented during the fall semester of 2013, and observations by the research team composed of a graduate student in mathematics and a mathematics professor were scheduled based upon instructor approval. Twenty-eight of the fifty-eight teachers agreed to participate in this initial study. Since some of these faculty members teach two completely different courses at the university, a total of thirty-six classroom observations occurred throughout the semester.

From the 36 classrooms participating in the study, 15 classes were taught by Graduate Teaching Assistants, 8 classes were taught by non-tenure faculty, and 13 classes were taught by tenured or tenure-track. There was a diverse amount of courses in this sample, ranging from college algebra to upper division mathematics. In the norm section of the results, the observed classes are grouped into five main categories: Precalculus, which includes college algebra to algebra with trigonometry courses; Applied Calculus which is a business calculus course; Calculus including Calculus I and II, differential equations, and computationally focused introductory linear algebra courses; Education which contains mathematics courses specifically designed for preservice elementary and secondary mathematics teachers; and Proof consisting of upper division proof-based mathematics courses. This study's observations included 11 Precalculus classes, 5 Applied Calculus classes, 11 Calculus classes, 4 Education classes, and 5 Proof classes. To determine the structure and reliability of the instrument, each class was observed once during the semester, and the analysis of the data collected from these thirty-six completed MCOP$^2$ forms is analyzed using exploratory factor analysis and classical text theory analysis.

## Results

The seventeen item MCOP$^2$ was analyzed using observations from thirty-six undergraduate mathematics classrooms. The researchers originally anticipated three factors to appear in the analysis with each factor corresponding to one of the three sections of the instrument (Student Engagement contains Items 1-5, Lesson Content contains Items 6-11, Classroom Culture and Discourse contains Items 12-17) (Gleason et al., 2014). However, as shown by the Scree Plot below (Figure 1), there are actually only one or two applicable factors. The third potential factor is not a legitimate component, despite its arguable location in the curvilinear region, since it is such a low eigenvalue. Furthermore, the Factor Matrix of a potential 3-Factor Model indicated the items from both Student Engagement and Classroom Culture and Discourse were actually loading onto the same factor, henceforth called Student Engagement and Classroom Discourse.

Figure 1: Scree Plot of Entire Protocol

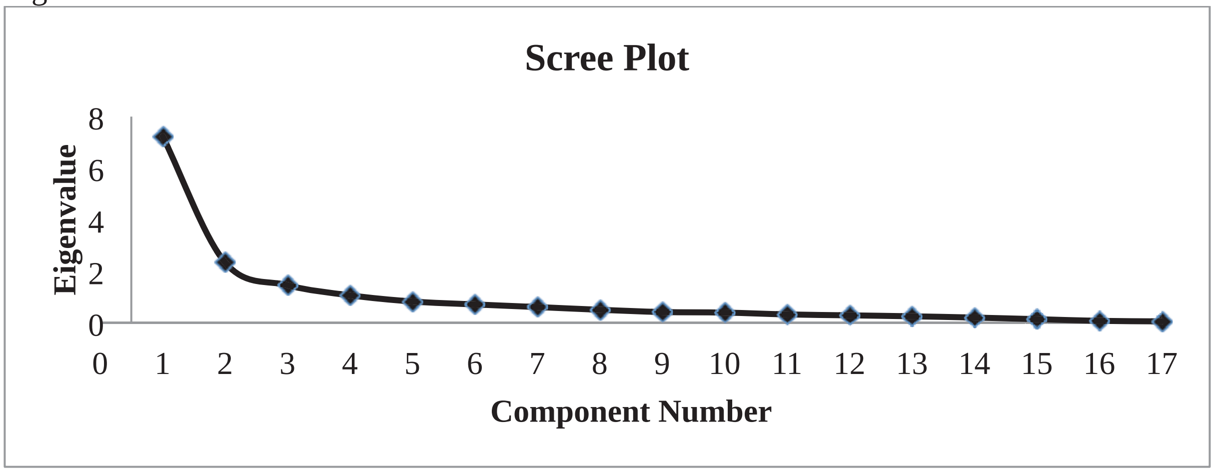
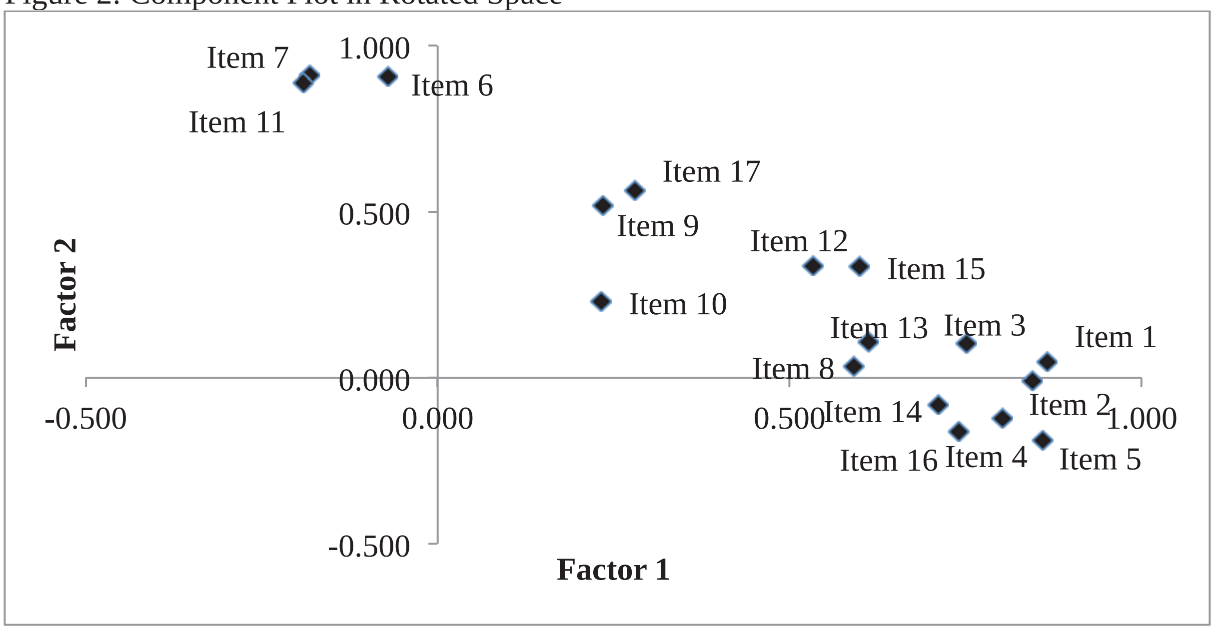



Figure 2: Component Plot in Rotated Space

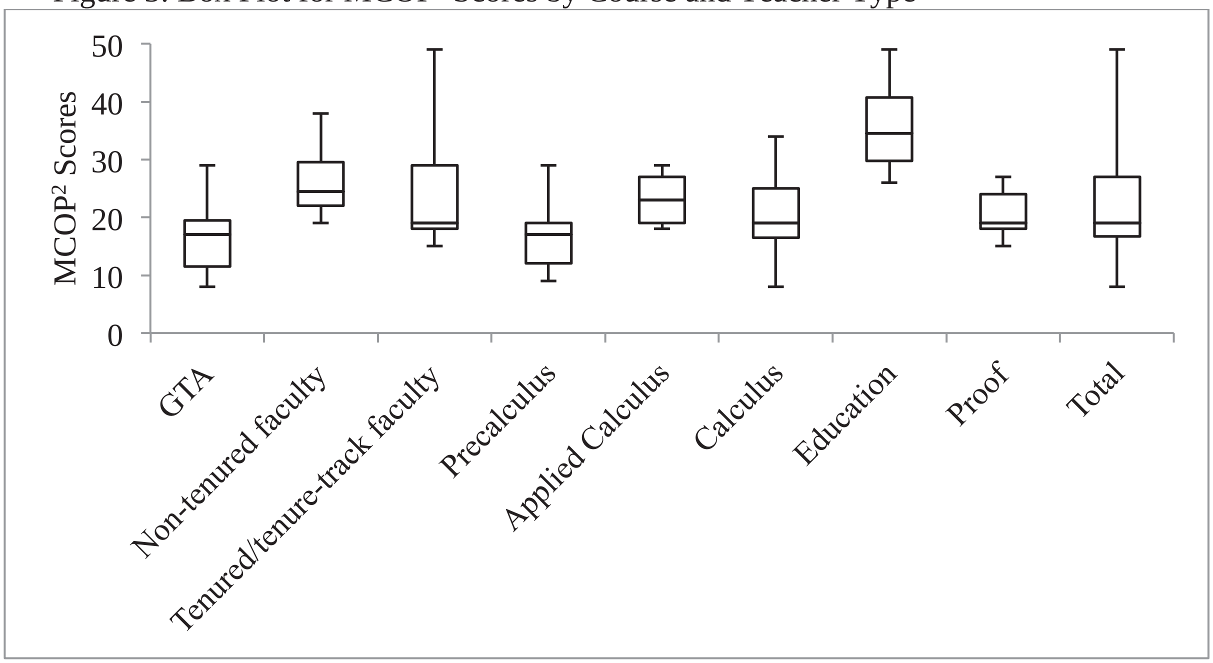



Solutions asking for two principle components to be extracted resulted in a 2-Factor Model explaining over 50% of the total variance. Using a promax rotation with Kaiser normalization, the component plot (Figure 2) indicates Items 1-5 and 12-16 are correlated, as expected, to Factor 1 (Student Engagement and Classroom Discourse) and Items 6, 7, 9, 10, and 11 are correlated to Factor 2 (Lesson Content). Items 8 and 17 did not load as expected, but instead loaded on the opposite factor. Since Item 8 was "The lesson promoted modeling in mathematics", it does not fit the theoretical construct of Student Engagement and Classroom Discourse, and so was removed from the reliability analysis of that subscale, but included on the overall instrument. Since Item 17 (The teacher uses student questions/comments to enhance mathematical understanding) fits well within the construct of lesson content, the results from this item are included in that subscale reliability, as well as the entire instrument.

The subscales of “Lesson Content” and “Student Engagement and Classroom Discourse” were both found to be unidimensional with over 50% of the variance contained in a single factor and so can be treated as subscales and analyzed for their own Cronbach’s alpha reliabilities of 0.779 and 0.907, respectively. In addition, if one considers the entire instrument to be unidimensional, then the entire protocol has a Cronbach’s alpha of 0.898. Therefore, the internal reliabilities are high enough for both subscales and the entire instrument to be used to measure at the group level, either multiple observations of a single classroom or single observations of multiple classrooms.

Norms from the 36 classroom sample used to create the factor analysis above are shown in Figure 3 to give future users of the $MCOP^2$ some standards of performance against which to assess the scores achieved by individuals or samples in their own data sets. In addition, from these norms, one can see that the instrument was able to differentiate between types of instructors and types of classes.

Figure 3: Box Plot for $MCOP^2$ Scores by Course and Teacher Type




## Conclusions

The overall instrument’s high coefficient alpha of .898 is noteworthy in that it demonstrates that the instrument is measuring something and is able to differentiate between classroom settings. Furthermore, the exploratory factor analysis indicates the $MCOP^2$gauges two main factors in an undergraduate mathematics classroom: “Lesson Content” and “Student Engagement and Classroom Discourse”. Thus when the instrument is separated into two sections, the $MCOP^2$’s Student Engagement portion demonstrates an exceptionally high level of internal reliability, proving the instrument successfully gauges an undergraduate mathematics classroom’s culture and student participation. Although not as remarkably high as the Student Engagement portion, the $MCOP^2$’s Lesson Content portion also shows high internal reliability, indicating the instrument also successfully measures the content of a mathematics lesson in a college level classroom.

This initial study has provided results indicating the Mathematics Classroom Observation Protocol for Practices ($MCOP^2$) is a reliable observational protocol for undergraduate mathematics classrooms. However, a much larger study of the instrument's reliability at the college level by testing the instrument in undergraduate mathematics classrooms at multiple higher education institutions, is needed as the data collected from observations at numerous community colleges, liberal arts schools, and other research universities would better examine and solidify the $MCOP^2$'s structure and reliability in a more general college mathematics classroom setting.

## References


AMATYC (1995). *Crossroads in Mathematics: Standards for Introductory College Mathematics before Calculus.* (D. Cohen, Ed.) Memphis, TN: American Mathematical Association of Two-Year Colleges. Retrieved from http://www.amatyc.org/?page=GuidelineCrossroads

AMATYC (2006). *Beyond Crossroads: Implementing Mathematics Standards in the First Two Years of College.* (R. Blair, Ed.) Memphis, TN: American Mathematical Association of Two-Year Colleges. Retrieved from http://beyonecrossroads.matyc.org/

APLU (2011). *The Common Core State Standards and Teacher Preparation: The Role of Higher Education*. APLU/SMTI, Paper 2. Washington, DC: Association of Public and Land-grant Universities. Retrieved from http://www.aplu.org/document.doc?id=3482

Barker, W., Bressoud, D., Epp, S., Ganter, S., Haver, B., & Pollatsek, H. (2004). *Undergraduate Programs and Courses in the Mathematical Sciences: CUPM Curriculum Guide, 2004*. Mathematical Association of America. 1529 Eighteenth Street NW, Washington, DC 20036-1358.

Berk, R. A., Naumann, P. L., & Appling, S. E. (2004). Beyond student ratings: Peer observation of classroom and clinical teaching. *International Journal of Nursing Education Scholarship*, *1*(1), 1-26.

California State University, East Bay. (2013). *Classroom observation worksheet.* Retrieved from http://www20.csueastbay.edu/faculty/ofd/resources/archive/class-worksheet-centra.pdf

Centra, J. A., Froh, R. C., Gray, P. J., & Lambert, L. M. (1976). A guide for evaluating teaching for promotion and tenure. Syracuse, NY: Syracuse University Center for Instructional Development.

Conley, D. T., Drummond, K. V., de Gonzalez, A., Rooseboom, J., & Stout, O. (2011). Reaching the Goal: The Applicability and Importance of the Common Core State Standards to College and Career Readiness. *Educational Policy Improvement Center*.

Gleason, J., Zelkowski, J., Livers, S., Dantzler, J., & Khalilian, J. (2014). Mathematics classroom observation protocol for practices: Validity and reliability. Preprint.

Harris, K-L., Farrell, K., Bell, M., Devlin, M., and James, R. (2008). Peer review of teaching in Australian higher education: A handbook to support institutions in developing effective policies and practices.

Hill, H. C., Charalambous, C. Y., Blazar, D., McGinn, D., Kraft, M. A., Beisiegel, M., Humez, A., Litke, E., & Lynch, K. (2012). Validating arguments for observational instruments: Attending to multiple sources of variation. *Educational Assessment*, *17*(2-3), 88-106.

King, J. E. (2011). Implementing the Common Core State Standards: An Action Agenda for Higher Education. *State Higher Education Executive Officers*.

Lawson, A. E., & Bloom, I., Falconer, K., Hestenes, D., Judson, E., Piburn, M.D., Sawada, D., Turley, J., Wyckoff, S. (2002). Evaluating college science and mathematics instruction. *Journal of College Science Teaching*, *31*(6), 388-393.

Marder, M., Abraham, L., Allen, K., Arora, P., Daniels, M., Dickinson, G., Ekberg, D., Walker, M. (2010). *The UTeach Observation Protocol (UTOP) training guide (adapted for video observation ratings)*. Austin, TX: UTeach Natural Sciences, University of Texas Austin.

NCTM. (2000). *Principles and Standards for School Mathematics*. Washington, D.C.: National Council of Teachers of Mathematics.

NGACBP & CCSSO. (2010). *Common Core State Standards for Mathematics.* Washington, DC: National Governors Association Center for Best Practices & Council of Chief State School Officers.

Sawada, D., Piburn, M., Falconer, K., Turley, J., Benford, R., & Bloom, I. (2000a). Reformed Teaching Observation Protocol (RTOP) (ACEPT Technical Report No. IN00-1). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.

Sawada, D., Piburn, M., Turley, J., Falconer, K., Benford, R., Bloom, I., & Judson, E. (2000b). Reformed teaching observation protocol (RTOP) reference manual. *Tempe, Arizona: Arizona Collaborative for Excellence in the Preparation of Teachers*.

Seldin, P. (1999). Current practices – good and bad – nationally. In P. Seldin & Associates (Eds.), Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions (pp. 1–24). Bolton, MA: Anker.

Spillane, J. P., Halverson, R., & Diamond, J. B. (2001). Investigating school leadership practice: A distributed perspective. *Educational researcher*, *30*(3), 23-28.

Tallahassee Community College. (2012). *Classroom observation worksheet.* Retrieved from https://www.tcc.fl.edu/FacultyStaff/CTLL/Documents/Classroom_Observation_Worksheet.pdf

University of New Mexico. (2006). *Classroom observation report.* Retrieved from http://www.unm.edu/~oset/SupportingDocuments/example%20classroom%20observation%20forms.pdf

Wainwright, C., Flick, L., & Morrell, P. (2003). The development of instruments for assessment of instructional practices in standards-based teaching. *Journal of Mathematics and Science: Collaborative Explorations*, *6*(1), 21-46.

Wainwright, C., Morrell, P. D., Flick, L., & Schepige, A. (2004). Observation of reform teaching in undergraduate level mathematics and science courses. *School Science and Mathematics*, *104*(7), 322-335.

Walkington, C., Arora, P., Ihorn, S., Gordon, J., Walker, M., Abraham, L., & Marder, M. (2012). Development of the UTeach observation protocol: A classroom observation instrument to evaluate mathematics and science teachers from the UTeach preparation program. Retreived from https://uteach.utexas.edu/sites/default/files/UTOP_Paper_Non_Anonymous_4_3_2011.pdf

Walkington, C., & Marder, M. (2013). Classroom Observation and Value-Added Models Give Complementary Information about Quality of Mathematics Teaching. Retrieved from https://uteach.utexas.edu/sites/default/files/WalkingtonMarderMET2013.pdf

Weiss, I., Pasley, J., Smith, P. S., Banilower, E., & Heck, D. (2003). Looking inside the classroom: a study of K-12 mathematics and science education in the United States. Horizon Research Inc., Chapel Hill, NC.

## Appendix A: MCOP$^2$ Items

1. Students engaged in exploration/investigation/problem solving.
2. Students used a variety of means (models, drawings, graphs, concrete materials, manipulatives, etc.) to represent concepts.
3. Students were engaged in mathematical activities.
4. Students critically assessed mathematical strategies.
5. Students persevered in problem solving.
6. The lesson involved fundamental concepts of the subject to promote relational/conceptual understanding.
7. The lesson promoted connections across the discipline of mathematics.
8. The lesson promoted modeling with mathematics.
9. The lesson provided opportunities to examine mathematical structure. (symbolic notation, patterns, generalizations, conjectures, etc.)
10. The lesson included tasks that have multiple paths to a solution or multiple solutions.
11. The lesson promoted precision of mathematical language.
12. The teacher's talk encouraged student thinking.
13. There were a high proportion of students talking related to mathematics.
14. There was a climate of respect for what others had to say.
15. In general, the teacher provided wait-time.
16. Students were involved in the communication of their ideas to others (peer-to-peer).
17. The teacher uses student questions/comments to enhance mathematical understanding.

## Appendix B: Sample MCOP$^2$ Item Descriptors

*5. Students persevered in problem solving.*

One of the Standards for Mathematical Practice (NGACBP & CCSSO, 2010) is that students will persevere in problem solving. Student perseverance in problem solving is also addressed in the Mathematical Association of America's Committee on the Undergraduate Program in Mathematics Curriculum Guide (Barker et al., 2004): "Every course should incorporate activities that will help all students…approach problem solving with a willingness to try multiple approaches, persist in the face of difficulties, assess the correctness of solutions, explore examples, pose questions, and devise and test conjectures."

Perseverance is more than just completion or compliance for an assignment. It should involve students overcoming a road block in the problem solving process.

| Score | Description |
|---|---|
| 3 | Students exhibited a strong amount of perseverance in problem solving. The majority of students looked for entry points and solution paths, monitored and evaluated progress, and changed course if necessary (NGA & CCSSM, 2010; Barker et al., 2004). When confronted with an obstacle (such as how to begin or what to do next), the majority of students continued to use resources (physical tools as well as mental reasoning) to continue to work on the problem. |

| | |
|---|---|
| 2 | Students exhibited some perseverance in problem solving. Half of students looked for entry points and solution paths, monitored and evaluated progress, and changed course if necessary (NGA & CCSSM, 2010; Barker et al., 2004). When confronted with an obstacle (such as how to begin or what to do next), half of students continued to use resources (physical tools as well as mental reasoning) to continue to work on the problem. |
| 1 | Students exhibited minimal perseverance in problem solving. At least one student but less than half of students looked for entry points and solution paths, monitored and evaluated progress, and changed course if necessary (NGA & CCSSM, 2010; Barker et al., 2004). When confronted with an obstacle (such as how to begin or what to do next), at least one student but less than half of students continued to use resources (physical tools as well as mental reasoning) to continue to work on the problem. There must be a road block to score 1 -3. |
| 0 | Students did not persevere in problem solving. This could be because there was no student problem solving in the lesson, or because when presented with a problem solving situation no students persevered. That is to say, all students either could not figure out how to get started on a problem, or when they confronted an obstacle in their strategy they stopped working. |

*7. The lesson promoted connections across the discipline of mathematics.*

This item focuses on helping students to see connections between different parts of mathematics. For early elementary grades, this could be a connection between measurement and counting or area models for multiplication. In the middle grades, this could be a connection between area and distributive property, or a connection between operations on different number systems. At the high school level an example would be connections between algebraic and geometric reasoning, or a connection between the different types of inverses. In an undergraduate classroom, this could be an opportunity for students to explore mathematical ideas from a variety of perspectives, or a connection to other subjects (both in and out of the mathematical sciences), or a connection to a contemporary topic from the mathematical sciences and its applications.

| Score | Description |
|---|---|
| 3 | Connections are emphasized throughout the lesson and/or are a major component of the lesson. |
| 2 | Connections are frequent throughout the lesson, but the connections are not a major component of the lesson. |
| 1 | A few connections are made in the lesson, but it is not frequent. |
| 0 | The lesson just makes no connections between mathematical topics. |