



**HAL**  
open science

# The structure of EQIPM, a video coding protocol to assess the quality of community college algebra instruction

Vilma Mesa, Irene Duranczyk, Laura Watkins

## ► To cite this version:

Vilma Mesa, Irene Duranczyk, Laura Watkins. The structure of EQIPM, a video coding protocol to assess the quality of community college algebra instruction. Eleventh Congress of the European Society for Research in Mathematics Education, Utrecht University, Feb 2019, Utrecht, Netherlands. <hal-02422677>

**HAL Id: hal-02422677**

**<https://hal.science/hal-02422677v1>**

Submitted on 22 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# The structure of EQIPM, a video coding protocol to assess the quality of community college algebra instruction

Vilma Mesa<sup>1</sup>, Irene Duranczyk<sup>2</sup>, Laura Watkins<sup>3</sup> and the AI@CC Research group

<sup>1</sup>University of Michigan, Ann Arbor MI United States; [vmesa@umich.edu](mailto:vmesa@umich.edu)

<sup>2</sup>University of Minnesota, Minneapolis MN, United States, [duran026@umn.edu](mailto:duran026@umn.edu)

<sup>3</sup>Glendale Community College, Glendale, AZ, United States, [laura.watkins@gccaz.edu](mailto:laura.watkins@gccaz.edu)

*Evaluating the Quality of Instruction in Post-secondary Mathematics (EQIPM) is a video coding instrument that provides indicators of the quality of instruction in community college algebra lessons (AI@CC Research Group, 2017)<sup>1</sup>. It grew out of two instruments that assess the quality of instruction in K-12 settings—the Mathematical Quality of Instruction (MQI) instrument (Hill, 2014) and the Quality of Instructional Practices in Algebra (QIPA) instrument (Litke, 2015). We present a revision of EQIPM prompted by results of factor analyses performed with preliminary data, then discuss next steps and research implications.*

*Keywords: Algebra, instruction, video coding, community colleges.*

Various reports have established an indirect connection between students leaving science, technology, engineering, and mathematics (STEM) majors because of their poor experiences in their STEM classes (Herzig, 2004; Rasmussen & Ellis, 2013). Most of these reports, however, are based on participants' descriptions of their experiences in the classes, rather than on evidence collected from large scale observations of classroom teaching (Seymour & Hewitt, 1997). When such observations have been made, they usually focus on superficial aspects of the interaction (e.g., how many questions instructors ask, how many students participate, or who is called to respond, Mesa, 2010) or their organization (e.g., time devoted to problems on the board, or lecturing, Hora & Ferrare, 2013; Mesa, Celis, & Lande, 2014). Undeniably, these are important aspects of instruction, yet these elements are insufficient to provide a characterization of such a complex activity as instruction.

A key concern in post-secondary mathematics education is the lack of preparation that mathematics instructors receive in their graduate education (Ellis, 2015; Grubb, 1999). We argue that the lack of a reliable and valid method to fully describe how instruction occurs hinders our understanding of the complexity of instructors' work in post-secondary settings and therefore limits the richness of preparation and professional development opportunities focused on the faculty-student-content

---

<sup>1</sup> The AI@CC Research group includes: Megan Breit-Goodwin, Anoka-Ramsey Community College; Anne Cawley, California State University, Pomona; April Ström, Chandler-Gilbert Community College; Randy Nichols, Delta College; Anna Bright, Patrick Kimani, Fern Van Vliet and Laura Watkins, Glendale Community College; Angeliki Mali, University of Groningen; David Tannor, Indiana Wesleyan University; Jon Oaks, Macomb Community College; Nicole Lang, North Hennepin Community College; Carla Stroud, and Judy Sutor, Scottsdale Community College; Saba Gerami, and Vilma Mesa, University of Michigan; Irene Duranczyk, Dexter Lim, and Nidhi Kohli, University of Minnesota. Colleges and authors are listed alphabetically.

interactions (Bryk, Gomez, Grunow, & LeMahieu, 2015). As part of a larger project that investigates the connection between the quality of instruction and student learning in community college algebra courses, we have developed an instrument, EQIPM (Evaluating Quality of Instruction in Postsecondary Mathematics), that seeks to characterize instruction. This paper focuses on the instrument and its validation, and to this effect we present the fourth revision prompted after results from exploratory and confirmatory factor analyses and a discussion with our advisory board suggesting several areas that needed refinement.

## **Theoretical Perspective**

We assume that teaching and learning are phenomena that occur among people enacting different roles—those of instructor or student—aided by resources of different types (e.g., classroom environment, technology, knowledge) and constrained by specific institutional requirements (e.g., covering preset mathematical content, having instructional periods of 50 minutes, see Chazan, Herbst, & Clark, 2016; Cohen, Raudenbush, & Ball, 2003). We focus on instruction, one of many activities that can be encompassed within teaching (Chazan et al., 2016) and define instruction as the interactions that occur between instructors and students with the mathematical content (Cohen et al., 2003). These interactions are influenced by the environment where they happen and can change over time. Empirical evidence from K-5 classrooms (5-11 year olds) indicates that ambitious instruction (Boston, 2012) is positively correlated with student performance on standardized tests (Hill, Rowan, & Ball, 2005). Understanding mathematics instruction requires attention to the disciplinary content and the mathematical knowledge for teaching and learning. Therefore, we assume, first, that quality instruction is illustrated through the experiences of instructors and students while interacting with mathematical content that have a significant impact on what students are ultimately able to demonstrate in terms of knowledge and understanding, and second, that it is possible to identify latent constructs that might account for the observed quality of instruction. We seek to corroborate findings from Hill et al. (2005) in the community college context.

EQIPM: Evaluating Quality of Instruction in Postsecondary Mathematics			
Features of Segment	Qual. Student-Content Interaction	Qual. Instructor-Content Interaction	Qual. Instructor-Student Interaction
Nature of Mathematics in Segment	Mathematical Reasoning and Sense-Making (S)	Making Sense of Mathematics (I)	Instructor-Student Continuum of Instruction
Modes of Instruction	Connecting across Representations (S)	Connecting across Representations (I)	Classroom Environment
Technology Use	Situating the Mathematics (S)	Situating the Mathematics (I)	Inquiry/Exploration
		Supporting Procedural Flexibility	Remediation of Student Errors and Difficulties
		Organization in the Presentation	
		Mathematical Explanations	
Mathematical Errors and Imprecisions in Content and Language			

**Figure 1: Dimensions and codes for the EQIPM 4.0 instrument**

Instruction is central to EQIPM. EQIPM (see Figure 1) was designed to assess the three main interactions that Cohen and colleagues (2003) used to define instruction: (1) Quality of Instructor-Student interaction, (2) Quality of Instructor-Content Interaction, and (3) Quality of Student-Content Interaction, which are hypothesized by our framework for instruction (Cohen et al., 2003, see Figure 1). An additional aspect, *Mathematical Errors and Imprecisions in*

*Content and Language*, is also hypothesized to be an indicator of the quality underlying all the interactions (Hill, personal communication, 2017). The *Mathematical Errors and Imprecisions in Content and Language* code is intended to capture all *uncorrected* errors or imprecisions that may occur within a segment; including those advanced by students. This code is the only code that attends to errors and imprecisions within the segment. Within each quality construct we propose codes adapted from existing rubrics (Hill, 2014; Litke, 2015) in order to capture specific observable behaviors that can be used as a proxy for the construct. Thus, the *Student Mathematical Reasoning and Sense-Making* code is a proxy for the quality of student interaction with content as it seeks to describe and qualify instances in which students make their thinking evident. Likewise, the codes *Connecting Across Representations* and *Situating the Mathematics* provide evidence for such interaction (Leinwand, 2014). In a similar way we chose four codes under the instructor content interaction, *Instructors Making Sense of Mathematics* a code that seeks to assess the extent to which instructors assist students in making sense of the mathematics they are teaching. This code is related, but different from, *Mathematical Explanations*, which attends primarily to the quality of the mathematical argumentation and justifications that instructors provide for any particular mathematical process or idea. While sense making might use informal language or everyday contexts, explanations require in addition that definitions or proofs be used in sound mathematical arguments. *Supporting Procedural Flexibility* was included in order to account for the quality of the teaching of procedures, an important component of the teaching of algebra (Litke, 2015; Star, 2005; Star & Newton, 2009). The last code in this dimension, *Organization in the Presentation*, originally from Litke's (2015) instrument, was included because there is evidence that such organization contributes to student performance (Cabrera, Colbeck, & Terenzini, 2001). Finally, the codes under the third dimension seek to capture the way in which instructors and students negotiate the work in the classroom, the extent to which their contributions are taken up to shape the direction of the lesson (*Continuum of Instruction*), how instructors respond to and seek to understand student misconceptions (*Remediation of Student Errors and Imprecisions*), and how the interactions produce a class environment that is conducive to learning (*Classroom Environment*). These codes directly reveal the quality of the student and instructor interactions. The codes under Features of Segment help characterize the activity in the segment without qualification (e.g., *Nature of math*: procedures, applications, etc.; *Modes of instruction*: lecture, group work, etc.; *Technology use*:

graphing calculators, etc.). These codes are not assessed on quality, only on presence and are used to describe the lessons.

## **Understanding how the instrument works**

In the Fall 2017 semester, we video-recorded 131 lessons in intermediate and college algebra classes from six different community colleges in three different states. These lessons ranged in duration between 45 and 150 minutes, and were taught by 40 different instructors (44 different unique courses video-recorded; 4 instructors taught 2 sections of a course). The majority of the lessons covered one of three topics: linear equations/functions, rational equations/functions, or exponential equations/functions. These topics were chosen because they offered opportunities to observe instruction on key mathematical concepts (e.g., transformations of functions, algebra of functions) and to attend to key ways of thinking about equations and functions (e.g., preservation of solutions after transformations, covariational reasoning), which are foundational algebraic ideas that support more advanced mathematical understanding (Breidenbach, Dubinsky, Hawks, & Nichols, 1992; Carlson, Jacobs, Coe, Larsen, & Hsu, 2002). The development of EQIPM was iterative process similar to that used by Hill and colleagues (2008) and by Litke (2015). Their instruments subdivided video-taped lessons into 7.5-minute segments, after testing the efficacy of 3 to 15-minute segments, they found that 7.5-minutes provided coders with a segment long enough to capture mathematical practices and events in a cohesive way when rating all segments within a lesson without compromising quality of coding that is experienced in long segments. In the Spring 2018 semester, our iterative process produced the EQIPM version 3a which was used in the first round of coding (AI@CC Research Group, 2017). Coders were given up to four consecutive 7.5-minute segments to reduce bias due to familiarity with the previous work. Coders recorded their ratings and justifications in a spreadsheet developed for this purpose after watching each segment. Each segment was rated on each of the codes on a 1 to 5 scale (see Mali et al., 2019 for more information on rating and exemplification of codes). Ten percent of segments were randomly chosen for double-coding. Each pair of coders held calibration meetings to discuss codes with ratings with a discrepancy greater than one point; the pairs reconciled their new rating by consensus.

Following an exploratory factor analysis [EFA] using 169 segments (Mesa, Duranczyk, Bardelli, & AI@CC Research Group, 2018) we identified a three-factor structure that corresponded to our hypothesized dimensions (see Figure 2a). This analysis also identified two problematic codes (*Remediation of Student Errors and Imprecisions* and *Mathematical Errors and Imprecisions in Content and Language*). A subsequent confirmatory factor analysis [CFA] with 306 coded segments, suggested instead a two-factor structure (Instructor-Content interaction and Student-Instructor interaction) and four problematic codes (one had poor loading, *Connecting Across Representations*, and three could not be used in the analysis, *Mathematical Errors and Imprecisions in Content and*

Exploratory Factor Analysis				Confirmatory Factor Analysis		
Codes	Factor 1	Factor 2	Factor 3	Codes	Factor 1	Factor 2
St. Mathematical Reasoning	<b>0.790*</b>	-0.001	-0.028	St. Mathematical Reasoning...	<b>0.844*</b>	
Instructor-Student Continuum	<b>1.055*</b>	-0.182	0.012	Instructor-Student Continuum...	<b>0.878*</b>	
Inquiry/Exploration	<b>0.594*</b>	0.039	-0.016	Inquiry/Exploration	<b>0.656*</b>	
Remediation Student Errors	<b>0.281*</b>	0.187	-0.057	Remediation of St. Errors...	<b>0.392*</b>	
Instr. Making Sense Procedures	-0.179	<b>0.661*</b>	0.022	Instr. Making Sense Proc.		<b>0.667*</b>
Supporting Proc. Flexibility	0.051	<b>0.307*</b>	<b>-0.283*</b>	Mathematical Explanations		<b>0.561*</b>
Organization Pres. of Procedures	0.072	<b>0.454*</b>	0.297	Situating the Mathematics		<b>0.480*</b>
Math. Err & Imprecision	0.023	<b>-0.224*</b>	-0.156	Supporting Procedural Flexibility		<b>0.409*</b>
Math. Explanations	-0.001	<b>0.954*</b>	<b>-0.792*</b>	Conn. across Representations		<b>0.289*</b>
Connect across Representations	0.036	-0.053	<b>0.300*</b>			
Situating the Mathematics	-0.05	0.222	<b>0.325*</b>			
Class Environment	0.374	0.011	<b>0.530*</b>			

The 3-factor solution had an adequate to good fit to the data:  $\chi^2 = 47.466$ ,  $p = 0.049$ , RMSEA = 0.051, 90% CI = [0.003, 0.082], CFI = 0.972, TLI = 0.944, SRMR = 0.065 (Hu & Bentler, 1999). Significant loadings at the 95% level. The three factors are weakly and positively correlated: corr Factor 1-Factor 2 = 0.184, corr Factor 1-Factor 3 = 0.245, and corr Factor 2-Factor 3 = 0.391. \* Significant loadings at the 95% level. a. Loading patterns suggest to exclude item in the EFA solution (Worthington & Whittaker, 2006). Factor 1: Quality of Instructor-Student Interaction. Factor 2: Quality of Instructor-Content Interaction; Factor 3: Quality of Student-Content Interaction.

The two-factor solution had a good fit to the data: RMSEA = 0.061, 90% CI = [0.039, 0.083], CFI = 0.958, TLI = 0.942, SRMR = 0.065. These two factors are weakly and positively correlated: corr Factor 1-Factor 2 = 0.351 ( $p < 0.001$ ) indicating that the two factors are correlated, but significantly distinguishable. \* Significant loadings at the 95% level. In CFA loadings are only estimated for the hypothesized factors. Factor 1: Quality of Instructor-Student Interaction; Factor 2: Quality of Instructor-Content Interaction.

**Figure 2: (a) Loadings for the EFA 3-factor solution; (b) Loadings for the CFA 2-factor solution**

*Language, Organization in the Presentation of the Procedure, and Classroom Environment*, see Figure 2b). In addition, in both analyses, some codes did not load into the hypothesized dimensions (e.g., *Situating the Mathematics, Connecting Across Representations, Student Mathematical Reasoning and Sense Making*). We called a meeting with our advisory board to discuss the conflicting results. Their review of the instrument and the findings helped us identify two issues that required substantial revision of the instrument (EQIPM 4.0, see Appendix for the revised definitions of the codes) to better describe the ratings and de-confound: 1) quality with quantifiable elements and 2) low quality with absence. We discuss these next.

### Confounding quality with quantifiable elements

For some of the codes we had assigned low or high ratings depending on either the time span in which the behavior was observed (or not observed) or other visible quantifiable aspects of the observed behavior. For example, for the code *Student Mathematical Reasoning and Sense Making* we were assigning a high rating to cases in which the work was “sustained” over the duration of the segment; if an instance was observed but it was “not sustained” or “not the focus” of instruction in the segment, the instrument directed coders to lower the rating. In another case, in the *Connecting Across Representations* code we accounted for number of representations that were visible in the segment in addition to the quality of the connections made; a high rating would be assigned only to lessons in which the instructor used 3 or more representations. Such approach resulted in lower ratings for segments in which instructors made really good connections but only used two representations; these segments were rated as a 3. We believe that this attending to number of representations possibly over quality of the connection might explain why this particular code had low loadings in both factor analyses. In EQIPM 4.0 we have eliminated language across all codes that refers to quantity or duration, and instead focused strictly on quality.

## **Confounding low quality with absence**

The rating system that we chose, a 1 through 5 scale, was adopted following prior work by Hill (2008) and Litke (2015). Our instrument included similar language in describing a rating of 1 which was assigned to two scenarios. For example, a 1 rating for *Mathematical Explanations* read as follows: “No mathematical explanations provided by instructor or student. **OR** Explanations do not include mathematical reasoning or justification; instructor or students provide only steps of a procedure.” This way of defining a 1 was problematic, because in some codes a rating of 1 meant not present (e.g., *Situating the Mathematics*), in others (e.g., *Classroom Environment*) it represented the lowest quality, and yet in others it meant both (e.g., *Mathematical Explanations*, the example above). Our advisory board pointed out that the ambiguity makes it difficult to interpret what a low score in a scale would represent. In EQIPM 4.0, the revision of the meaning of a rating of 1 also required a decision about the scale for assessing quality. We had followed Litke’s practice of using a rating of 3 to be “modal practice,” meaning practices that could be expected in many instances or that would not be extraordinary (in terms of both high and low quality). Ratings of 2 and 4 were defined as “in between” levels, less than or better than 3, without being the extremes. Defining the scale in this way was problematic for several codes, especially for those in which 1 could mean not present. In those codes the change from 1 to 2 was not commensurable to a change from 2 to 3, 3 to 4, or 4 to 5 on our rating scale. The recommendation was to use a four-point scale in order to force the coders into differentiating four levels of quality, with 1 being the lowest and 4 being the highest. We assigned a rating of 0 for “not present” which will facilitate their exclusion as needed.

## **The revised EQIPM 4.0 instrument**

The current version of EQIPM is presented in Figure 1. Once the data are recoded we will run a split sample EFA/CFA to understand the revised instrument’s factor structure. The preliminary analyses with the previous instrument lend support to the structure underlying the instrument. One issue we face is the high number of segments in which lecture is the main mode of instruction (75%, 688 of 920 segments coded as of this writing had lecture combined with other modes of instruction, e.g., group work, and 487 of those used only lecture). We believe that with the revised instrument we will be able to differentiate interactions that are solely led by the student. If there are enough segments in which students lead the interactions, we will run an analysis to see whether the three-factor structure holds.

As we move forward we will account for (1) the multi-level structure of our data, specifically, segments within lessons, lessons within instructors, and instructors within colleges, which will address the possible non-independence of some of these codes, (2) lesson duration (e.g., in some lessons, student work individually at the end) and (3) coder bias. We anticipate that the structure will be stable in these cases as well and that as lecture diminishes as consequence of exposure to professional development, we may see that data will fit with three distinct factors as student interaction with content increases.

## **Acknowledgment**

Funding for this work was provided by the National Science Foundation award EHR #1561436. Any opinions, findings, and conclusions or recommendations expressed in this material are those of

the author(s) and do not necessarily reflect the views of the National Science Foundation. We thank Emanuele Bardelli and Inah Ko for analytical assistance and Saba Gerami for comments to earlier versions of the proposal. We thank the faculty who participated in the project. Without them this work would not have been possible.

## References

- AI@CC Research Group. (2017). *Evaluating the Quality of Instruction in Postsecondary Mathematics. Instrument for video coding*. Glendale Community College, Scottsdale Community College, University of Michigan, University of Minnesota. Phoenix, AZ; Scottsdale, AZ; Ann Arbor, MI; Minneapolis, MN.
- Boston, M. (2012). Assessing instructional quality in mathematics. *Elementary School Journal*, 113(1), 76-104.
- Breidenbach, D., Dubinsky, E., Hawks, J., & Nichols, D. (1992). Development of the process conception of function. *Educational Studies in Mathematics*, 23, 247-285.
- Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Cambridge, MA: Harvard Education Press.
- Cabrera, A. F., Colbeck, C. L., & Terenzini, P. T. (2001). Developing performance indicators for assessing classroom teaching practices and student learning. *Research in Higher Education*, 42(3), 327-352.
- Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education*, 33, 352-378.
- Chazan, D., Herbst, P., & Clark, L. (2016). Research on the teaching of mathematics: A call to theorize the role of society and schooling in mathematics instruction. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of research on teaching* (5th ed., pp. 1039-1097). Washington DC: American Educational Research Association.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25, 119-142.
- Ellis, J. (2015). Three models of graduate student teaching preparation and development. In D. M. Bressoud, V. Mesa, & C. L. Rasmussen (Eds.), *Insights and recommendations from the MAA National Study of College Calculus* (pp. 117-122). Washington, DC.
- Grubb, N. W. (1999). *Honored but invisible: An inside look at teaching in community colleges*. New York: Routledge.
- Herzig, A. (2004). Becoming mathematicians: Women and students of color choosing and leaving doctoral mathematics. *Review of Educational Research*, 74, 171-214.
- Hill, H. C. (2014). *Mathematical Quality of Instruction (MQI)*. Retrieved from Harvard, MA: <https://hu.sharepoint.com/sites/GSE-CEPR/MQI-Training/Shared%20Documents/MQI%204-Point.pdf>
- Hill, H. C., Blunk, M., Charalambous, C., Lewis, J., Phelps, G., Sleep, L., & Ball, D. L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430-511.

- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42, 371-406.
- Hora, M. T., & Ferrare, J. J. (2013). Instructional systems of practice: A multidimensional analysis of math and science undergraduate course planning and classroom teaching. *Journal of the Learning Sciences*, 22(2), 212-257.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Leinwand, S. (2014). *Principles to action: Ensuring mathematical success for all*. Reston, VA: National Council of Teachers of Mathematics.
- Litke, E. G. (2015). *The state of the gate: A description of instructional practice in algebra in five urban districts*, (PhD). Harvard University, Cambridge, MA.
- Mali, A., Cawley, A., Duranczyk, I, Mesa, V., Ström, A., & Watkins, L. (2019). Identifying sense-making in algebra instruction at U.S. post-secondary colleges. In U.T. Jankvist, M. Van den Heuvel-Panhuizen, & M. Veldhuis (Eds.). *Proceeding of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. XXX-YYY). Utrecht, the Netherlands: Freudenthal Group & Freudenthal Institute, Utrecht University.
- Mesa, V. (2010). Student participation in mathematics lessons taught by seven successful community college instructors. *Adults Learning Mathematics*, 5, 64-88.
- Mesa, V., Celis, S., & Lande, E. (2014). Teaching approaches of community college mathematics faculty: Do they relate to classroom practices? *American Educational Research Journal*, 51, 117-151.
- Mesa, V., Duranczyk, I., Bardelli, E., & AI@CC Research Group. (2018). *An exploratory factor analysis of EQIPM, a video coding protocol to assess the quality of community college algebra instruction*. Ann Arbor MI.
- Rasmussen, C. L., & Ellis, J. (2013). Who is switching out of calculus and why. In A. M. Lindmeier & A. Heinze (Eds.), *37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 73-80). Kiel, Germany.
- Seymour, E., & Hewitt, N. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Westview Press: Boulder CO.
- Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education*, 36, 404-411.
- Star, J. R., & Newton, K. J. (2009). The nature and development of experts' strategy flexibility for solving equations. *ZDM*, 41(5), 557-567.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806-838.

#### Appendix: Revised Definition of the Codes

Code	Revised Definition: The following codes assess...
<b>Student Mathematical Reasoning and Sense-Making:</b>	student utterances that showcase reasoning and sense-making about mathematical ideas (e.g. drawing logical conclusions, providing conjectures, counter-claims, reasoning and engaging cognitively in problem solving)
<b>Connecting Across</b>	the connections that instructors or students make between and across representations

<b>Representations:</b>	of the same mathematical problems, ideas, and concepts
<b>Situating the Mathematics:</b>	how instructors or students make connections to other aspects of the algebra curriculum, related topics, or the broader domain of mathematics, situating and motivating the current area under study within a broader context.
<b>Instructors Making Sense of Mathematics:</b>	how instructors attend to specific aspects of mathematics (e.g., solution, symbols, conditions) to clarify their nature
<b>Mathematical Explanations:</b>	how mathematical reasons and justification for why something is done are provided
<b>Supporting Procedural Flexibility:</b>	how instructors identify what procedure can be applied, and when and where to apply them, or makes connections across procedures
<b>Organization in the Presentation:</b>	how complete, detailed, and organized the instructor's or students' presentation (either verbal or written) of content is when outlining or describing the mathematics at hand.
<b>Instructor-Student Continuum of Instruction:</b>	how the investment that students make in their own learning and development of mathematical understanding by expressing thoughtful ideas that advance their learning.
<b>Classroom Environment:</b>	how instructor and students create a respectful and open environment in their classroom conducive to learning; high quality mathematical work is the norm.
<b>Inquiry / Exploration:</b>	the amount of exploration and inquiry of the mathematics that students do in the classroom
<b>Remediation of Student Errors and Difficulties:</b>	how remediation in which student misconceptions and difficulties with the content are addressed
<b>Mathematical Errors and Imprecisions in Content or Language:</b>	mathematically incorrect or problematic use of mathematical ideas, language, or notation.