EXAMINING HOW TEACHERS ENACT THE SUGGESTIONS OF A COACH: CRITIQUE OF A METHODOLOGY

Ryan Gillespie
University of Idaho
gill6762@vandals.uidaho.edu

Julie M. Amador
University of Idaho
jamador@uidaho.edu

Jeffrey Choppin
University of Rochester
jchoppin@warner.rochester.edu

In this methodology paper, we present a methodology for characterizing how teachers use coaches' suggestions. We identified suggestions from planning conversations and explored the extent to which the teachers implemented the suggestions in enacted lessons. The planning conversations took place within online content-focused coaching cycles. A primary challenge confronting content-focused coaches when working one-on-one with teachers is finding a productive balance between giving suggestions and inquiring into teachers' practices through reflective questioning. This paper articulates a process for identifying suggestions made by a coach during a planning conversation and an analytic process for examining how a teacher takes up the suggestion during lesson implementation. We discuss the methodological challenges we encountered and tradeoffs in our decisions related to low- and high-inference claims.

Keywords: Inservice Teacher Education/Professional Development, Research Methods, Coaching

Coaching as a form of professional development is a promising practice (Campbell & Griffin, 2017; Ellington, Whitenack, & Edwards, 2017). Within mathematics education, content-focused coaching (e.g. West & Cameron, 2013) is a common model. Content-focused coaching involves iterative cycles in which a coach works one-on-one with a teacher, with a focus on students' mathematical learning goals. Each coaching cycle contains three sequential components: a pre-conference discussion to plan a lesson; a collaboratively taught lesson; and a post-conference discussion to debrief the lesson (Bengo, 2016; West & Staub, 2003).

Research on coaching has highlighted two competing stances for how coaches talk with teachers: *reflective* or *directive* (Deussen, Coskie, Robinson, & Autio, 2007; Ippolito, 2010; Sailors & Price, 2015). Coaches using a *reflective* stance emphasize collaborative inquiry in which the coach elicits ideas from the teacher; these ideas become the basis of the coach-teacher discussion (Ippolito, 2010). Coaching moves associated with a reflective stance include probing questions and low-inference, non-evaluative observations as means to catalyze teacher thinking (Costa & Garmston, 2016). In contrast, a *directive* coaching stance involves the use of suggestions and evaluative feedback (Ippolito, 2010). The challenge in content-focused coaching is to find the right balance between when to provide a teacher with direct assistance in the form of a suggestion and when to employ an inquiry stance (West & Staub, 2003). It is crucial for researchers within mathematics education to explore these stances and their impact on teacher learning and uptake of new practices.

Despite the importance of mathematics coaches strategically choosing appropriate actions when working with teachers, little is known about how mathematics coaches using a content-focused coaching model interact with teachers (Gibbons & Cobb, 2016). Furthermore, little research exists on how direct assistance from a coach during a coaching cycle supports a teacher to implement a lesson. As a first step in addressing this gap, this paper outlines a methodology for analyzing how teachers take up a coach's suggestions when planning and enacting lessons. Specifically, this study is guided by the question: How do we characterize the extent to which a teacher uses the suggestions of a coach during the implementation of a lesson? This methodology would help mathematics educators better understand the impact of a coach's suggestions on the pedagogical actions of a teacher and serve as a first step towards the larger inquiry of how the discursive actions of a coach impact teachers' practices.

In: Sacristán, A.I., Cortés-Zavala, J.C. & Ruiz-Arias, P.M. (Eds.). (2020). *Mathematics Education Across Cultures: Proceedings of the 42nd Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Mexico. Cinvestav / AMIUTEM / PME-NA. https:/doi.org/10.51272/pmena.42.2020

Methodological Processes Applied to Coaching Conversations

To identify the suggestions made by the coach during the collaborative planning conversations, we used results from a broader analysis of the coaching conversations. These results came from a broader study that analyzed planning and debriefing conversations within content-focused coaching cycles from four coaches paired with eight teachers over a period of two years. In that broader analysis, we focused on the discursive moves of mathematics coaches during coaching cycles. We developed a codebook to analyze the transcripts of planning and debriefing discussions between coaches and teachers; this codebook characterized the discursive moves of the coaches and teachers as well as the content of the conversations. The section of the codebook that focused on the discursive moves of the coach was comprised of five broad categories; including suggestions (see Figure 1). We defined a *suggestion* as a statement from the coach recommending an action for the teacher.

Coaching Stance	Discursive Move	Description	Example
Reflective	Invitation	Statement or question that invites the teacher to reflect or respond	"What might be some strategies we could use to increase student participation?"
	Description	Statements that share a direct observation and do not contain inference, interpretation, judgement or opinion	"I noticed that during the turn-and-talk, Alex did not say anything to his partner."
Directive	Suggestion	Statement that recommends an action	"I think we should use a turn-and-talk prior to the whole class discussion."
	Explanation	Statements that provide an interpretation or rationale of an event, interaction, or mathematical idea	"Turn-and-talk is a powerful strategy to use prior to a whole class discussion because it increases student participation."
	Evaluation	Statements that offer praise or critique	"I think it was a great idea to use a turn-and-talk."

Figure 1: Excerpt from the larger codebook focusing on coaching discursive moves

We parsed the transcripts of the planning and debriefing conversations into stanzas, which included a coach's statement and the participant's response, as well as text needed for context (Saldaña, 2013). This broader data set included the analysis of n = 1719 stanzas from 41 transcripts of coaching conversations. We coded stanzas in pairwise teams after a lengthy calibration process that involved five researchers. We met via video conferencing software, Zoom, to reconcile disagreements. Kappas ranged from 0.39 to 0.65, considered moderate to strong reliability (Landis & Koch, 1977).

The following is an excerpt from a coach's comments that was coded as a *suggestion*:

> One of the really nice moves you can do if the group shares a thought about something, and it's somewhat ambiguous, is you can turn to the class and say, "Can someone else use their own words to explain what Dave is saying?"

In this comment, the coach recommended the teacher prompt students to paraphrase a peer's explanation as a means to increase student participation in classroom discussions.

Piloted Version of Data Analysis Process for Coaches' Suggestions

The purpose of this paper is to detail a methodological process for identifying suggestions made by coaches and characterizing how those suggestions were taken up by the teachers during lesson implementation. We wanted to understand how teachers incorporated the suggestions coaches provided during planning meetings into their teaching. We describe our current analytic attempts to

highlight the methodological affordances and challenges of the work and to gain insight from others in the mathematics education community. Knowing the extent to which teachers follow coaches' suggestions in their teaching is important information for coaches as they plan to support teachers.

Coding Suggestions from the Coaching Transcripts

We collected all stanzas previously identified as involving a coach's suggestion. If a stanza contained multiple distinct suggestions (i.e. more than one action was recommended by the coach), each suggestion was placed into a different row in the spreadsheet where we tracked the suggestions, with the goal of distilling a coach's suggestions to the smallest granular size. This process converted the unit of analysis from a stanza to an individual suggestion. For example, a coach made the following statement to the teacher during a planning conversation:

> Whereas, if you really want them to be able to understand the formula, you're going to be asking different questions about, where did this come from? What do we know about volume? What does volume mean? Those kinds of more probing questions as they're working or as they're thinking about it, and as you're launching. Then, what I'm also thinking about is your ticket out the door idea. This idea of, do you want to do some checking in with students in terms of their understanding about volume related to the cylinder and the cone, either before they leave you Monday, or possibly Tuesday, so that you get a sense of, beyond just your questioning and asking each individual group, would you want to have—would it be helpful to have some kind of—some documentation to look back at in terms of students' understanding? I'm just thinking of a little, mini half-sheet, or something, if it would be, again, helpful to figure out where kids are in their thinking.

We coded this statement as two distinct suggestions. First, the coach suggested that the teacher ask different questions about the concept of volume. Then, the coach provided suggestions about the ticket out of the door. These two suggestions were listed separately.

Next, two coders individually coded each of the coach's suggestions for the content of the suggestion using the codebook (see Figure 2).

Content						
Lesson Goals			Contextualizing			
Instructional Practice	Math Content	Coaching	School/Classroom Context	Math Content or Curriculum	Students(s)	
Lesson Content						
Introduction/Launch	Task Description & Sequencing	Questions/Questioning	Summary	Assessment	Adaptions	Discourse
Lesson Features						
Instructional Needs	Misconceptions/ Challenges	Anticipating	Selecting/Sequencing	Connecting	Participation Structure	Next Steps	Student Strategy
Technology						
Technical Logistics/Challenges	Referencing Swivl/Annotations/Video	Use of Technology				

Figure 2: Content codebook, with categories in purple and codes in yellow.

Each coder created a concise statement capturing (a) the *general nature of the suggestion* and (b) the *specific action in the suggestion* recommended by the coach (see Figure 3). The coders then met to reconcile and reach consensus for their content codes and created statements. For example, during a planning conversation, a coach said:

> Kids are always surprised that getting one of each happens so much more often than getting two of the same. Your experimental data is definitely going to show that. What if you asked them, "If you toss a coin twice, what are the things that could happen?" You ask them that at the beginning, just to get an idea of where they are.

We applied the content codes of *question*/*questioning*, *assessment*, and *introduction*/*launch* because the coach recommended the teacher ask a question during the launch phase of the lesson to assess student thinking prior to beginning a task (see Figure 3). The general nature of the suggestion was written as: "Ask students questions during launch to assess their understanding." The specific action suggested by the coach was captured by the coders through the statement, "Ask the question 'If you toss a coin twice, what are the things that could happen?' during the launch to assess student understanding."

Stanza Excerpt	Content Codes	General Nature of the Suggestion	Specific Action in the Suggestion
Kids are always surprised that getting one of each happens so much more often than getting two of the same. Your experimental data is definitely going to show that. What if you asked them if you toss a coin twice, what are the things that could happen? You ask them that at the beginning, just to get an idea of where they are.	Assessment, Question/ Questioning, Introduction/ Launch	Ask students questions during launch to assess their understanding	Ask the question "If you toss a coin twice, what are the things that could happen?" during the launch to assess student understanding.

Figure 3: Excerpt from the coding spreadsheet

If there was not a specific action within the suggestion, the *specific action in the suggestion* column was left blank. Additionally, if the suggestion was deemed to not be observable within the lesson video, we coded the suggestion as *not observable*. For example, if the coach suggested an action for the teacher in future lessons beyond the current coaching cycle, the suggestion would be *not observable* during the lesson video.

Coding Enactments of Suggestions

The coders independently watched the video of the implemented lesson to identify how the teacher followed the general and specific suggestions from the coach. For each suggestion marked in the spreadsheet generated from analyzing the coaching transcripts, the coders worked on two levels. First, they considered if the general nature of the suggestion was *present* or *not present* in the lesson. Second, they considered the extent to which the specific action of the suggestion was taken up by the teacher. To code the specific action, they chose between the following codes that represented a continuum of uptake: *not present*, *partially adhere*, *mostly adhere*, or *fully adhere* (see Table 1).

Table 1: Coding Scheme for a Teacher's Enactment of a Specific Suggestion

Code	Description	Example (connected to Figure 2)
Not present	The teacher did not enact any part of the coach's suggestion during the lesson.	The teacher did not ask any questions during the lesson launch.
Partially present	The teacher enacted only a single part of the coach's suggestion during the lesson.	The teacher asked a question during the lesson launch but the question does not relate to assessing understanding.

Mostly adhere	The teacher enacted multiple, but not all, parts of the coach's suggestion during the lesson.	The teacher asked a question during the lesson launch to assess understanding but used a question worded differently the question suggested by the coach.
Fully adhere	The teacher enacted all parts of the coach's suggestion during the lesson.	The teacher asked the exact question suggested by the coach during the lesson launch.

In cases where the teacher did not have the opportunity to enact a suggestion from the coach, the instance was coded as *no opportunity*. For example, if the coach suggested the use of a specific teacher talk move during a whole-class summary discussion of a task (e.g. Smith & Stein, 2011) but the class period ended before the teacher was able to begin the summary discussion, we coded that as the teacher not having the opportunity to enact the coach's suggestion. The code *no opportunity* is different than *not observable* in that *no opportunity* was based on a suggestion that could be observable but was beyond the scope or inclusion of the implemented lesson due to constraints such as time or classroom happenings in the moment of the lesson. Figure 4 contains a flowchart summarizing this analysis process.

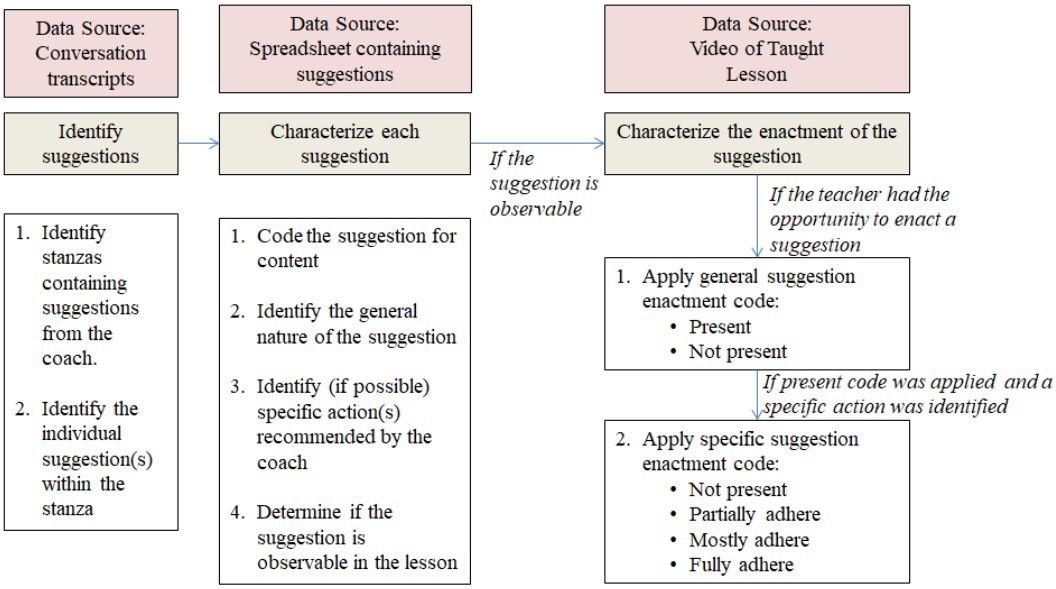

Figure 4: Flowchart of analysis process

Methodological Affordances and Challenges

In this section we describe the methodological affordances of our analysis process as well as the challenges we encountered when: a) identifying suggestions given by the coach within conversation transcripts; b) categorizing a coach's suggestions based on how the suggestion invited action from the teacher; and c) characterizing the extent to which teachers took up a suggestion when teaching a lesson. We discuss these affordances and challenges within each step of our analysis process to highlight methodological issues when characterizing direct assistance from coaches and the impact of that direct assistance on teachers' instructional practices.

Identifying Suggestions

The first challenge we encountered was establishing a reliable coding process for determining when coaches made suggestions. The coaches used a range of discursive moves when recommending an action to a teacher (Gillespie, Amador, & Choppin, 2019), making the identification of suggestions problematic. For example, a coach said:

> Another idea for your launch is, again, I don't know how I feel about this, but is to give them a pool with side length five, and actually have them work on their—or do an independent time of—and explaining the tiling. You've got this patio going around the outside of the five by five square and how many tiles would it take for that patio.

This statement was coded as a *suggestion* because the coach recommended a task and participation structure for the teacher to use during the introduction of a lesson. A similar call for action was also expressed by a coach through the question, "What if you did just one example, not of a one by one, but a five by five or something?" Even though this discursive move is a question, the recommendation of modifying an activity is embedded within the question. Thus, this statement was also coded as a *suggestion*. We coded discursive moves as *suggestions* if the move contained an explicit recommendation for action or a clear, implicit recommendation embedded within a conversational move. This decision allowed the coding process to capture a wider variety of suggestive moves from coaches. However, because this decision required the researchers to analyze conversational moves for the underlying intent, the coding entailed some inference of intent. Besides acknowledging the inferential nature of some of the coding, we have struggled with how to methodologically characterize the nature of our inferences. In summary, accurately identifying suggestions from a coach requires analyzing discourse for both explicit and implicit recommendations directed at the teacher; analyzing for implicit intent, however, causes the coding process to become increasingly inferential.

Characterizing Suggestions

The second challenge in developing the analytic process involved characterizing the actionable nature of the suggestions. Initially, we realized that many of the suggestions were broadly stated, providing considerable latitude in how teachers might interpret and use them. For example, during a planning conversation, a coach said, "You probably want to look for students that have done the different strategies so that you know who you want to share how they did it, how they organized it." In this instance, the coach recommended the teacher monitor student thinking in order to select students to share different strategies during a summary discussion, but did not offer specific guidance as to which strategies to select. In the absence of a specific action, we coded the suggestion as a general suggestion and then coded for whether or not that general action was *present* or *not present* during the teaching of the lesson. However, the general nature of the suggestion made it more difficult to consider the tangible impact of the coach's suggestions. We found other suggestions to be specific; these suggestions provided exact language describing an action a teacher should use during the lesson. For example, a coach said, "You could even ask a question about—there's eight options there. 'If you play it 40 times, how many times would you expect so-and-so to win, and how many times would you expect the other guy to win?'" In this example, the coach recommended the teacher ask questions and provided two specific questions for the teacher to use. Suggestions that were specific were easier to identify in the enacted lesson and code but also represented a more localized impact on practice. Our coding scheme allowed us to distinguish between general and specific suggestions, but we were able to make more definitive decisions for specific suggestions. This poses a dilemma: our coding of uptake may only capture highly tangible evidence of impact while missing potentially more powerful and broader impact from the more general suggestions. Our challenge can be summarized in the following way: general suggestions have the potential to convey broader pedagogical principles than more specific ones, but entail greater inference when coding.

Characterizing Take Up of Suggestions

A third challenge was determining the extent to which a teacher enacted the precise details in a specific coach suggestion. We used a continuum to code the enactments. Coding along a continuum for the presence of specific details in a specific suggestion entailed qualitative characterization of the suggestion. Our continuum had these four characterizations, from low to high presence of a suggestion: a) no part of the suggestion, b) a single part of the suggestion, c) multiple parts of the suggestion, or d) the suggestion in its entirety. As a result, we created the codes *not present*, *partially adhere*, *mostly adhere*, or *fully adhere* to use when coding the teacher's enactment of specific suggestions (i.e. Table 1). As an example of coding a teacher's enactment of a suggestion using this continuum, during a planning conversation a coach said:

> They could even try tossing it three times. You could make it into a situation where a couple kids are playing a game. One kid wins if all three of the tosses match. One kid wins if only two of them match, and is it a fair game idea?

This was coded as a specific suggestion because the coach recommended the teacher use an activity and provided specific context and questions to use in the activity. During the lesson, the teacher facilitated an activity in which students flipped three coins but did not frame the activity as a game that may or may not be fair. Because only a single part of the specific suggestion was enacted by the teacher, the suggestion was coded as *partially present*. This qualitative characterization allowed us to describe accurately the extent to which the teacher enacted the suggestion but posed challenges related to the reliability of our coding.

To limit inference and potentially support more reliable coding, our coding continuum did not evaluate the effectiveness of the enacted suggestion; if a specific suggestion was evident in the lesson, we did not then consider whether or not the enactment was productive. So, while the teacher carried out the suggestion, the way that happened, and the implications of what happened subsequently in the lesson, may not have been aligned with the coach's intent. Avoiding evaluating the productiveness of the enactment of a suggestion has implications for the ways we consider the impact of a coach's suggestion on teacher practice, but using a low-inference coding scheme had its advantages. Although the process was anecdotally deemed reliable by coders during our first round of coding, additional calibration will be needed to reliably code the enactment of specific suggestions using the four leveled codes. In summary, characterizing the ways teachers enact a specific coach suggestion requires the use of a continuum, posing challenges for reliability. This challenge can be partially mitigated by limiting focus to the ways the teacher followed the explicit language of the coach's instruction, not considering intent or effectiveness; yet this limits the extent to which the data can be used to make claims about the impact of the coach on productive practice.

Discussion

Coaching is an increasingly popular professional development practice in mathematics education; however, more needs to be known about how coaches interact with teachers (Gibbons & Cobb, 2016). We presented the methodological challenges and opportunities in characterizing how the *suggestions* of a coach impact the practice of a teacher. Focusing on coaches' suggestions is one start to more fully understanding how coaching influences teaching practice. Researchers should continue to examine how the actions of a coach impact the practice of a teacher. More specifically, because coaches use discourse as a primary tool to engage teachers (Costa & Garmston, 2016; Heineke, 2013), we anticipate similar challenges exist in identifying coaching discursive moves beyond suggestions (e.g. invitational or evaluative). Knowing how discursive moves (i.e. Figure 1, Deussen, Coskie, Robinson, & Autio, 2007; Ippolito, 2010; Sailors & Price, 2015) impact the development of teachers would be beneficial for knowing how to support teachers through coaching. Thus, the specific challenges discussed within this paper relate to more global challenges that will be

encountered during any analysis of how the discursive actions of a coach impact the practice of a teacher.

By illuminating the challenges and opportunities learned through our work, we aim to support future researchers by emphasizing the complexity of analyzing the relationship between the discursive moves of a coach and a teacher's practice. For example, when identifying a coach's suggestions within varied discursive moves in a full coaching conversation, we had to consider language that explicitly and implicitly communicated a recommendation to capture the nuanced ways the coach provided suggestions. Coding is this way required us to develop rules that reliably connected the language of a coach to their intent to provide a suggestion. These coding decisions introduced inference. However, not considering intent through implicit language would have constrained our ability to identify instances in which coaches leveraged their expertise to share actionable ideas with teachers for use in a lesson. We anticipate similar challenges will be encountered in identifying other discursive moves of coaches.

The second challenge was characterizing the suggestions in ways that considered the implications on a teacher's enactment of the suggestion. We created two classifications, *general nature* and *specific action*, but found general suggestions can promote broader pedagogical principles than more specific ones, but require greater inference when coding. We posit this challenge will also apply to coding schemes attempting to capture the impact of other coaching discursive moves. For example, a coach using broad or general invitational moves may provide the teacher with more latitude to reflect, but analysis may require higher levels of inference.

Our third challenge was limiting inference when coding for the enactment of specific suggestions; accomplished in part by not considering the effectiveness of the enacted suggestion. Low-inference coding in this context has advantages but disregarding effective enactment prohibits the study from claims about how the actions of a coach connects to changes in the quality of a teacher's practice. In other words, our current process will allow us to make low-inference claims about how teachers used the suggestions of the coach but prevent us from making claims about how the suggestions influenced the quality of the lesson. Discussion of these challenges invites future mathematics education researchers to consider methods to overcome these obstacles.

Acknowledgments

The material is based upon work supported by the National Science Foundation under Grant #1620911.

References

Bengo, P. (2016). Secondary mathematics coaching: The components of effective mathematics coaching and implications. *Teaching and Teacher Education, 60*, 88-96.

Campbell, P. F., & Griffin, M. J. (2017). Reflections on the promise and complexity of mathematics coaching. *The Journal of Mathematical Behavior, 46*, 163-176.

Costa, A. L., & Garmston, R. J. (2016). *Cognitive coaching: Developing self-directed leaders and learners*. Lanham, MD: Rowan & Littlefield.

Deussen, T., Coskie, T., Robinson, L., & Autio, E. (2007). *"Coach" can mean many things: five categories of literacy coaches in Reading First* (Issues & Answers Report, REL 2007–No. 005). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. Retrieved from http://ies.ed.gov/ncee/edlabs.

Ellington, A., Whitenack, J., & Edwards, D. (2017). Effectively coaching middle school teachers: A case for teacher and student learning. *The Journal of Mathematical Behavior, 46*, 177-195.

Gillespie, R., Amador, J., & Choppin, J. (2019). *Exploring the discursive differences of mathematics coaches within online coaching cycle conversations*. In S. Otten, A. G. Candela, Z. de Araujo, C. Haines, & C. Munter (Eds.),

Proceedings of the forty-first annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (pp. 442-450). St. Louis, MO: University of Missouri.

Gibbons, L. K., & Cobb, P. (2016). Content-focused coaching: Five key practices. *The Elementary School Journal, 117*(2), 237-260.

Heineke, S. F. (2013). Coaching discourse supporting teachers' professional learning. *The Elementary School Journal, 113*(3), 409-433.

Ippolito, J. (2010). Three ways that literacy coaches balance responsive and directive relationships with teachers. *The Elementary School Journal, 111*(1), 164-190.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

Sailors, M., & Price, L. (2015). Support for the improvement of practices through intensive coaching (SIPIC): A model of coaching for improving reading instruction and reading achievement. *Teaching and Teacher Education, 45*, 115-127.

Saldaña, J. (2013). *The coding manual for qualitative researchers*. Los Angeles: SAGE Publications.

Smith, M. S., & Stein, M. K. (2011). *5 practices for orchestrating productive mathematical discussions*. Reston, VA: National Council of Teachers of Mathematics.

West, L., & Staub, F. C. (2003). *Content-focused coaching: Transforming mathematics lessons*. Portsmouth, NH: Heinemann.

West, L., & Cameron, A. (2013). *Agents of change: How content coaching transforms teaching and learning*. Portsmouth, NH: Heinemann.