Continuing Evolution of Research on Teaching and Learning: Exploring Emerging Methods for Unpacking Research on Teachers, Teaching, and Learning

Chandra Hawley Orrill, Zarina Gearty, and Kun Wang

1 Introduction

In 1987, Medley offered an explanation of the "present state of the art of research in teaching" (p. 105). By doing so, he outlined the categories of variables that could be studied and provided strong guidance for high-quality research on these variables. His guidance suggested that research should seek to find out why teaching quality varies widely and to do that, one must have a conceptualization of what good teaching is, an instrument that is valid for distinguishing good teaching from poor teaching, and a plan for collecting accurate data and for analyzing that data. While aspects of research on the relationship between all of the variables that shape the complex act of teaching students in a formal learning environment have remained unchanged since 1987, much has changed in educational research including the emergence of new methodologies and the increasing presence of technology both as a tool for teaching and learning and as a tool for research, making research on the connections between and among the variables both richer and more flexible than Medley originally suggested.

C. H. Orrill (✉)
Kaput Center for Research and Innovation in STEM Education, University of Massachusetts Dartmouth, 506 Langley St, Fall River, MA 02720, USA
e-mail: corrill@umassd.edu

Z. Gearty
Kaput Center for Research and Innovation in STEM Education, University of Massachusetts Dartmouth, 3 Rumford Road, Norton, MA 02766, USA
e-mail: zmadan@umassd.edu

K. Wang
Kaput Center for Research and Innovation in STEM Education, University of Massachusetts Dartmouth, 120 Johnson Road, Winchester, MA 01890, USA
e-mail: kwang1@umassd.edu

A. Manizade et al. (eds.), *The Evolution of Research on Teaching Mathematics*, Mathematics Education in the Digital Era 22,
https://doi.org/10.1007/978-3-031-31193-2_12

In this chapter, our goal is to provide some clear examples of the ways in which the field of mathematics education has been able to pursue the interactions put forth by Medley (1987) thanks to the emergence of new theories, research methods, and technologies. We discuss some of the critical ways in which the research landscape has changed since the original article was published. First, we look at methodologies that have become more widespread since Medley introduced his framework. Second, we explore some examples of research methods that offer new ways of thinking about research questions related to teaching and learning in formal environments. These methods are important because they have opened opportunities to look across the variables in ways that were unavailable to researchers until recently. We finish by considering the opportunities that have been created by technology. These have changed the ways we collect data and the data we collect. We close by briefly discussing how the changes discussed in this chapter specifically relate to presage-process–product research (Medley, 1987). Our goal in this chapter is not to present a comprehensive review of the literature, rather, we seek to highlight both where the field is now in terms of research methods and tools as well as to provide examples of the ways in which Medley's framework is being pursued in newer research. We have chosen to rely on Fig. 3 in Chap. 1 (Manizade et al., 2023) rather than the original Medley model for the purposes of our discussion, except where noted.

2 Changing Methodologies

Since Medley's framework was originally proposed, the field has seen the emergence of new research methodologies (i.e., scientific frameworks), methods (i.e., specific approaches), and tools (i.e., instruments) that allow innovative lenses with which to make sense of the multi-faceted enterprise that is teaching and learning. In this chapter, we offer brief overviews of just a few of the methods currently available for answering questions related to the presage-process–product model. These tools sometimes allow us ways to look at connections between more than two variable categories (e.g., cultural-historical activity theory) or allow us to conceive of research as a web of interconnected studies all serving to develop a larger theory (e.g., Design-Based Research). Below, we talk about the emergence of Qualitative and Mixed Methods methodologies as well as emerging psychometric models, then we introduce several methods that provide new ways of thinking about the interconnected nature of the variables for teaching and learning. These include teaching experiments, Design-Based Research, cultural-historical activity theory (CHAT), and quantitative ethnography.

3 Qualitative and Mixed Methods Research Methodologies

In 1987, qualitative research was not often used in education. While there were certainly some examples of qualitative research emerging (e.g., Erlwanger, 1973), those studies were not as widely accepted as quantitative studies. However, with the shifts in ontology and increased acceptance of new methods, the field of mathematics education research became more open to—and, indeed, primarily focused on—qualitative research. The reason is simple: quantitative research methodologies are particularly appropriate for a model of teaching and learning that relies on transfer of knowledge from the expert to the novice. As we moved toward theories of learning that were more grounded in constructivism, socio-cultural theories, and critical theories, new questions were being asked. As an array of new learning theories emerged, the definitions of teaching and learning became more diverse and even questions could be pursued. Rather than asking if teachers who took teaching methods courses in mathematics or science are more effective than those who did not, researchers began wondering in what ways particular backgrounds might shape learning experiences (Type E—teacher's competencies, knowledge and skills—and Type F—pre-existing teacher characteristics) and their interaction on Type A (student learning outcome) variables (e.g., Manizade et al., 2023; Medley, 1987)) and how teachers conceive of making content learnable for children (Type D variables interacting with Type B and A variables (e.g., Manizade et al., 2023; Medley, 1987)). Further, with the emergence of new theories of learning, the definitions of what constitutes learning, and, thus, how learning is measured, also changed. Consider, for example, Wenger's theory of situated cognition (Wenger, 1998). Within this theory, "learning" is defined as a change in participation, because learning is viewed as becoming a member of a community of practice. Thus, as one learns, one becomes a fully-integrated member of the community of practice. If participation is the goal, a written assessment of content knowledge is no longer an appropriate instrument for measuring learning and new approaches need to be developed. Thus, it is consistent with the rise of the cognitive, socio-cultural, and critical theories that qualitative research would become a critical tool for research.

Once qualitative methodologies were established as a norm within the field, it was natural for some researchers to use qualitative and quantitative methods together to better understand a phenomenon. Thus, mixed methods approaches have been used by some researchers to understand the interactions inherent in the learning environment. Grounded in pragmatism (Johnson et al., 2007), mixed methods research is a methodology that combines qualitative and quantitative methods to produce "defensible and usable research findings" (p. 129). For example, a researcher may conduct a survey (quantitative), then conduct interviews with a subset of participants (qualitative) to more deeply understand the findings of the survey (Creswell, 2014). For example, in one recent study (Starrett et al., 2021), researchers used surveys and interviews with teachers and students to understand how teacher's proximity to their school impacted their use of place-based education, thus using mixed methods to connect Type I (external context) variables to Type D (pre- and post-active mathematics teacher activities), Type C (interactive teacher activities), and Type B (student learning activities). Another approach to mixed methods analysis is to use the various analyses to dig into specific aspects of the data. For example, an approach the first author of this chapter has used (e.g., Izsák et al., 2010; Orrill & Cohen, 2016), included mixture Rasch analysis of an assessment of teacher's math knowledge. Our goals was to identify specific mathematics tasks that teachers found difficult or with which different groups of teachers had different experience. From those data, we were able to identify specific items on which to focus in the qualitative analysis of the interviews. Using this approach, we were able to not only see how teachers performed on the assessment, but also to generate assertions about why they performed in these ways, thus providing us with additional information for designing effective instruction. These methods allowed us to more thoroughly understanding Type E variables.

For understanding the ways in which teachers, their experiences, and their actions intersect with student learning, access to a wide array of methodologies is crucial. While quantitative research is still used, studies using qualitative and mixed methods approaches are as accepted in most venues as quantitative research. The critical factor in high-quality research is not the methodologies and methods used, but rather the alignment of the methods and methodologies to the research questions.

4 Continuing to Develop Quantitatively: Emerging Psychometric Models

Methods for conducting quantitative research have also continued to develop since 1987. While quantitative research has remained theoretically grounded in positivism and still adheres to the methodological frameworks that were in place in the 1980s, quantitative research has benefitted tremendously from increased access to computers and the emergence of new models that has been possible because of computers. Now that nearly everyone has access to extremely powerful computing technology, quantitative analysis can be more robust and more accessible than ever. Particularly important for presage-process–product research are the myriad statistical and psychometric models that have emerged in the past few decades. In this section, we briefly introduce four such models that have played a role in our own research on teacher knowledge and student learning: Item Response Theory (IRT), mixture Rasch Models, Diagnostic Classification Models, and Topic Models. We have selected these four models because each provides researchers with different information about learning. Further, we included IRT because of its widespread use. Each of the four offers a way to better connect the variables highlighted by Medley. We do not, however, intend this as an exhaustive list.

Item Response Theory (Baker, 2001; Baker & Kim, 2004) is probably the most influential psychometric model in widespread use as it has largely replaced classic test theory for scoring standardized tests. Rather than simply assigning a score indicating how many items are correct or incorrect, IRT provides various kinds of information. First, IRT provides a score that expresses a participant's performance in terms of the number of standard deviations above or below the mean of participant scores. The second piece of data provided by IRT is information regarding the difficulty of an item, where difficulty is reported as the probability that someone scoring at the identified difficulty level would have a 50% chance of answering the item correctly. IRT provides researchers with information that allows them to consider not just whether "learning" has occurred, but where there may be deficits in aspects of content knowledge as well as a relative weighting of participant's performances. Because of the information that it can return, IRT is currently a critical component in the development of learning trajectories (e.g., Clements et al., 2011; Confrey et al., 2017), which sit at the intersection of student learning (Type A), teacher planning (Type D), and teacher knowledge of students (Types G & H).

Building from IRT, mixture Rasch models (Izsák & Templin, 2016; Rost, 1990) look for latent trends in the patterns of responses among participants to determine whether all the participants should be placed along the same continuum or whether there are groups within the data that performed in ways different from others. This approach has been used to identify patterns of reasoning among teachers. These patterns highlight that performance on an assessment can be tied to patterns in teacher's reasoning about the content (e.g., Izsák et al., 2010; Orrill & Cohen, 2016). The data can also be used to capture a different kind of "learning". Rather than focusing only on whether participants have improved their scores on an assessment, researchers can also determine whether participants have changed latent classes. Such change would indicate a fundamental shift in the ways the participants are reasoning about the mathematics items on the assessment (Izsák et al., 2010). While mixture Rasch has primarily been used for in-depth consideration of Type E variables, we assert that it could be used as a lens for understanding the relationship between Type E and Types C and D (how teachers plan and implement instruction). It could also be readily used to look at connections between Type A and Type C and D variables if an assessment were given to students and correlated to observations of classroom practice.

Another emerging family of models is Diagnostic Classification Models (DCMs; Bradshaw et al., 2014; Izsák & Templin, 2016; Rupp et al., 2010). DCMs require an a priori defining of the specific attributes each item of an assessment measures (e.g., Tatsuoka et al., 2016). From that mapping, analysis is done on participant's performance, and results are reported as the probability that the participant has mastered each individual attribute. For example, in Bradshaw et al. (2014), the authors identified four attributes being measured by the assessment of fractions: referent unit understanding, partitioning and iterating, appropriateness, and multiplicative comparison. The attribute inventories that are returned in place of traditional test scores can provide insights into specific aspects of understanding demonstrated by a given sample, thus providing data that can shape the instructional experiences for participants. As with mixture Rasch models, DCMs provide an opportunity to connect teacher or student understanding of mathematics (Type A or E) to the activity in the classroom (Type B or C).

One final emergent psychometric model is Topic Modeling (e.g., Blei, 2012). Topic Models allow a statistical analysis of qualitative data to show a change in patterns of language usage. Topic models rely on looking for particular words in natural text or natural speech to see their patterns of co-occurrence. From those patterns, groupings are created that separate participants. For example, in Kim et al. (2017), the researchers found three main themes in their analysis of student's responses on a science assessment. The first theme featured answers that included appropriate technical terms for middle grades science students (e.g., change, variable, dependent). The second theme was discipline-specific terms (e.g., energy, population, kinetic). The third theme focused on everyday language (e.g., put, stronger, big, think). Across four assessments, the participating middle school students shifted from using the everyday language topic to the other two topics (Linking Type B to Type C variables). These results suggested that students were learning about the discipline. Topic Modeling is particularly important for measuring learning through a socio-cultural lens.

New psychometric models do not fundamentally change the design and limitations of quantitative research. Thus, they have some limitations outlined by Medley in terms of looking at relationships between the variables. However, the new models allow measurement of different kinds of learning (e.g., change in participation or in natural language use, rather than acquisition of knowledge), and they open opportunities for mixed methods approaches such as those described in the discussion of mixed methods above. So, even within the realm of quantitative research, there are more tools available to support asking questions in new ways and looking at relationships through different theoretical lenses than was possible in 1987.

5 Emerging Research Approaches

In this section, we introduce four of the approaches to research that have changed the ways in which we can answer questions about teaching and learning in formal (and informal) contexts. As with the discussion above, we do not assert that these are the only approaches appropriate for research in the presage-process–product framework. Rather, these are tools that have been used, or are emerging in use, in mathematics education and the learning sciences to answer research questions related to the variables in Medley's framework and the connections between them. Certainly, there are many other approaches that could also be used for this purpose. Below, we discuss: Teacher Experiments, Design-Based Research, Cultural-Historical Activity Theory (CHAT), and Quantitative Ethnography. Teaching experiments are featured because of their prominence in mathematics education research over the past three decades, while the other approaches were selected because they offer robust and diverse pathways for making sense of the complexity of learning environments through their analytical lenses or through iterative implementation. For each, we describe what it is, some benefits and limitations for the approach, and some examples of studies done with the approach.

6 Teaching Experiments

6.1 What is It?

Teaching experiments stem from Piaget's clinical interviews (Steffe & Thompson, 2000) and have roots in Russian education research (e.g., Davydov, 1975). Teaching experiments are fundamentally constructivist and have been used with a wide variety of constructivist perspectives, ranging from radical constructivism to social constructivism (Cobb, 2000). In a teaching experiment, the researcher serves in the role of a teacher and conducts a series of teaching episodes, usually working with a small group of students or one individual (Cobb & Steffe, 1983). The key goal is to develop a "living model of student's mathematics" (Steffe & Thompson, 2000, p. 284), testing and revising instructional activities designed to support student learning (McClain, 2002).

Teaching experiments go beyond the scope of a clinical interview by aiming to help the researcher understand the change and progress of a student's mind rather than just the current state of the mind. The teacher-researcher constructs a conjecture about student's mathematical knowledge, then tests the conjecture with teaching episodes designed to move the student's understanding forward, reformulating the conjecture after each episode (Cobb & Steffe, 1983; Steffe & Thompson, 2000). Initial hypotheses can be abandoned based on student's responses as it is vital that the teacher-researcher allows the student's contributions to guide the trajectory of the teaching episode.

Typically, a teaching experiment consists of the teacher-researcher, the student(s), and an observer. The role of the observer is to witness and document student's reaction and behavior. The teacher-researcher, constantly interacting with the students and instantaneously reacting to the students, may not be able to capture all relevant observations (Cobb, 2000). The teacher-researcher engages the students in instructional tasks or activities to observe and promote mathematical learning and reasoning by posing tasks and asking follow-up questions (Steffe & Thompson, 2000). The data that is collected is qualitative and meant to record the models of student's mathematical understanding. The teacher-researcher uses this data to revise conjectures as teaching episodes progress and to revise the activities in the episodes. Ultimately, a model of student thinking about the specific content or topic is generated by the researcher, with a variety of qualitative data to support the model (Cobb & Steffe, 1983). From the perspective of the framework of research on teaching mathematics adapted from Medley (Manizade et al., 2023), teacher experiments allow conjectures to be made about how student mathematics learning outcomes (Type A) are shaped by the interactive mathematics teacher activities (Type C) that are pre-active activities (Type D) from considering student's engagement in previous mathematics learning activities (Type B). In a sense, while the outcome of the teaching experiment is a theory about student's learning, the experiment itself is an iteration of the relationships of Type A, B, C, and D variables conducted by a person with competency, knowledge, and skill (Type E) in the mathematics, in student learning, and in designing instructional interventions.

6.2 Benefits and Limitations

One clear benefit of teaching experiments is the insight they provide about how to support a student or small group of students to move forward in their understanding of specific concepts. A unique feature of teaching experiments is that the researcher is directly involved with the teaching. Therefore, the researcher should have teaching experience and the ability to interact and engage with students (Steffe & Thompson, 2000). The goal is to elicit and support thinking during these interactions; teacher-researchers should be cognizant of how their actions and language are perceived by students (Tallman & Weber, 2015).

Teaching experiments are powerful tools for understanding learning, however, they are also very challenging. One clear challenge of this method, particularly for inexperienced researchers, is that data collection and data analysis are simultaneous during the series of episodes (Tallman & Weber, 2015). Additionally, to demonstrate the evolution of the conjectures and model building, the researcher-teacher needs to maintain on-going documentation of the reasoning for decisions and the interpretation of student's thinking (McClain, 2002). Self-reflexivity becomes a key assumption, where the researcher-teacher acknowledges that he/she is an active participant of the student's constructions (Steffe & Thompson, 2000; Tallman & Weber, 2015). A common data analysis method after the series of teaching experiments is retrospective analysis, changing and revising the hypothesized model (Cobb & Steffe, 1983).

6.3 Examples

Many researchers used teaching experiments as exploratory tools, usually as part of larger projects. Simon and colleagues (2018) conducted a teaching experiment with a single student as part of the Measurement Approach to Rational Number (MARN) Project. Their goal was to understand how instruction could promote student's construction of the concept of multiplication with whole number and fractions, and to develop a hypothetical learning trajectory based on their analysis. The participant was a fifth-grade student, Kylie, that the research group had been working with for two years, conducting various clinical interviews and teaching experiments. One of the teaching experiments involved the use of a computer application called Java Bars as the instructional tool. Simon served as the teacher-researcher and posed several multiplication tasks designed to explore Kylie's changing conceptions about the meaning of the multiplier. The research group initially used the concept of generalizing assimilation, as defined by Piaget (1952) as the theoretical base for their hypothetical learning trajectory. They hypothesized that the progression of their instructional tasks would stimulate changes in the assimilatory structure of the student. However, during the initial teaching episodes, the researchers were not seeing any evidence of conceptual change and modified the teaching tasks. Using retrospective analysis of the data collected, Simon and colleagues revised their hypothetical learning trajectory conjecture to rather be stimulated by reflective abstraction, also a construct defined by Piaget (1952). When reporting their findings, the authors included a detailed description of the progression and rationale of the learning trajectory. The research group went on to conduct more teaching experiments with the revised instructional tasks and hypothetical learning trajectory (Simon et al., 2018). This research is focused on the interaction of student mathematics learning activities and student mathematical outcomes (Type B and Type A)—that is, how does an instructional intervention affect learning. Because it was a teaching experiment, though, it extended to Type C, interactive mathematics teacher activities, because one of the outcomes of this work was a hypothetical learning trajectory which could be used to guide other teachers for supporting student learning. Finally, consistent with the Framework of research on teaching mathematics as shown in Fig. 3 of Manizade and colleagues (Manizade et al., 2023), Type G research (individual student characteristics, abilities, and personal qualities) is also tacitly happening as the teacher-researcher is consistently assessing the student's abilities and understandings to make the instructional decisions that lead to particular learning activities.

Teaching experiments can also be done with larger groups of students. A study conducted with 299 undergraduate calculus students by Wagner and colleagues (2017) consisted of eight teaching episodes which were designed to study the change in student's ability to generate examples for the purpose of understanding novel concepts. The researchers formulated a hypothesized learning sequence and developed an instructional sequence of tasks and questions. The learning progression was broken down into intended student's awareness and behavior on specific skills and views. Over the course of eight teaching episodes, the teacher-researcher introduced the tasks, which progressed from more rigid to more open-ended to allow students to show their views and ability to generate examples. Data analysis was done using emergent codes from the participant's words and phrases from the reflections and written tasks. The evidence showed a progression of positive changes in the student's views of generating examples. The researchers revised their proposed learning sequence for the third iteration of the teaching experiment, where new students were chosen who had not yet learned calculus material, to test their revisions.

As with the Simon et al. (2018) example, this research approach follows a Type C-B-A flow, moving from considering interactive mathematics teacher activities to student mathematics learning activities to student mathematics learning outcomes. The researcher plays the role of a teacher, so data can be gathered on Type C, interactive mathematics teacher activities. Second, the research group designs a set of activities based on data from Type G (individual students characteristics, abilities, and personal qualities) and Type A (student learning outcomes) to simulate student learning experiences that could occur in the classroom, allowing the study of Type B (student mathematics learning activities) variables. Finally, the researchers analyze data and make assertions on how the activities impact student learning. Both groups used iterative design to develop a learning theory, ultimately laying out a trajectory of activities for teacher to implement and student to experience (Type C and Type B). Simon et al. (2018) implicitly also study the individual student Kylie, providing details about her abilities and personal qualities (Type G). In summary, teaching experiments are an extension of clinical interviews and align to constructivist learning theories. The researcher becomes a teacher in this methodology where he/she formulates, tests, and revises a hypothesis of a model for a change in student thinking as a response to some instructional sequence (Cobb, 2000; Steffe & Thompson, 2000).

7 Design-Based Research

7.1 What is It?

Design-Based Research (DBR) approaches originated from a desire to pursue research questions that cannot be answered in a laboratory setting (e.g., Brown, 1992; Collins, 1992). Underlying the development of DBR was a desire to develop an approach that overcame the issues in attempting to apply results from education laboratory studies into actual classrooms (Cobb et al., 2003; McKenney & Reeves, 2013). Over time, Brown's and Collin's notions of "design experiments" matured into an approach known by many names, that we refer to as Design-Based Research (DBR; Design-Based Research Collective (DBRC), 2003). DBR focuses on the development and refinement of theory along with the development and refinement of innovations that embody that theory (e.g., Barab & Squire, 2004; Cobb et al., 2003; DBRC, 2003; McKenney & Reeves, 2013). DBR is an approach to research that relies on a trajectory of inter-connected studies conducted, often over several years, rather than a single study (Cobb et al., 2003). It is inherently grounded in partnerships between researchers and practitioners.

A unique feature of DBR is the dual goal of generating a theory and developing and refining a particular intervention that embodies that theory (McKenney & Reeves, 2013; Sandoval, 2014). Researchers focus on both problematizing the context and on using theories to generate usable knowledge (DBRC, 2003). The development of such theories is a key component of DBR (Cobb et al., 2003). Through the iterative processes, conjectures are made, tested in the natural setting, revised based on the outcome, and tested again. The theory becomes emergent through this process and is refined at the end of the project (Barab & Squire, 2004). Because of this focus on developing theory and innovation together, DBR projects tend to include serious consideration of student learning outcomes (Type A), student mathematics learning activities (Type B), interactive mathematics teacher activities (Type C) and pre or post-active mathematics teacher activities (Type D). Many also include consideration of teacher competencies, knowledge, and skills (Type E); external context variables (Type I); and teacher development and experiences (Type J).

DBR projects usually span several years. The reason for this is twofold. First, DBR focuses on iterative in design (Cobb et al., 2003; DBRC, 2003), which involves multiple implementations, data collection and analysis cycles. Second, to understand an innovation and the theory it embodies in a way that moves toward generalizable knowledge, different grain sizes must be considered. Studies may focus on a single tool with a few students, then that tool used in a classroom, then that tool used in the context of the delivery of a piece of curriculum, etc. Because the studies focus on a series of related questions of different grain sizes, they often benefit from mixed methods approaches across the lifespan of the research (DBRC, 2003). Rather than having confined control variables, multiple dependent variables such as classroom environment and learning outcomes are examined to generate a deep understanding of the issues and the effect of the intervention (Cobb et al., 2003).

In DBR, researchers partner with various stakeholders to achieve the goals of refining theory and refining the intervention. Interventions can be educational products, policies, or programs (McKenney & Reeves, 2013). Examples of stakeholders are teachers, school leaders, coaches, and subject matter experts. These participants become an integral part of the development and implementation of the design, sharing their expertise to collaboratively work through the project (DBRC, 2003). Much of the work is conducted in a natural authentic setting, such as schools and classrooms; the context is problematized and studied as a vital part of understanding the learning and teaching that occurs (Barab & Squire, 2004).

The overall structure of DBR is flexible and iterative, but also systematic. It is a sequence of approaches rather than just one approach (Barab & Squire, 2004). Several models of approaches have been offered by researchers (e.g., Eljersbo et al., 2008; McKenney & Reeves, 2012; Reeves, 2011). Most of these models include the initial phase of exploring and analyzing a problem, followed by the construction of a design and then reflection and evaluation. Since the entire process is iterative and usually non-linear, most researchers using DBR work back and forth through those phases. Theories are developed and tested throughout the process until enough evidence and data is gathered for a mature theory and usable knowledge. Usable knowledge can be declarative knowledge, such as describing a certain phenomenon or prescriptive knowledge, such as ways to facilitate learning with a certain intervention (McKenney & Reeves, 2019). In the initial phase, researchers study a setting and develop testable conjectures about how to address an educational problem or how to influence a change in students learning (Cobb et al., 2003). Data analysis becomes an ongoing process as both researchers and practitioners aim to deepen their understanding of phenomenon that occur in the natural setting (Barab & Squire, 2004). This collective partnership and iterative design process can be seen as unique features of DBR whose purpose is to close the gap between educational research and classroom practice and to further theoretical knowledge that can influence change in settings facing similar problems (DBRC, 2003; McKenney & Reeves, 2019). The DBR approach has been applied in various sectors of education such as learning sciences, curriculum development, instructional design and teacher professional development (See special issues of journals such as *Educational Researcher (*2003, 31(1)), *Journal of the Learning Science* (2004, 13(1)) and *Educational Psychologist* (2004, 39(4)).

7.2 Benefits and Limitations

One of the key benefits of DBR is its ability to inform the development of a usable intervention while also yielding a generalizable theory. This two-faceted benefit ensures that both the immediate outcome of the project (the intervention) has educational merit while also ensuring that there is something beyond a single application of the theory that can support teaching and learning. This ensures that the theory can continue to inform practice beyond the lifespan of the intervention.

Due to its multi-faceted design, DBR can be a challenging approach even for experienced researchers. The role of the researcher is less defined and more fluid; she can be the designer and the implementor, which can introduce threats to validity and objectivity (Barab & Squire, 2004). Furthermore, the researcher needs to anticipate and communicate means of support for the various groups of people involved in the project, who often can have different opinions and perspectives on educational issues (Cobb et al., 2003). Time and personal commitment are devoted to maintaining close and respectful relationship with partnerships (Cobb et al., 2003). Another source of difficulty arises from the various sources of data and the extended period of collection time (DBRC, 2003). Various techniques for data collection and analysis are often required along with an appropriate balance between rich data and a surplus of data to ensure validity (McKenney & Reeves, 2013) and often retrospective analysis is needed for theory development (Cobb et al., 2003). Despite these limitations and challenges, researchers have found DBR to be useful for a wide range of studies. A few such studies are highlighted in the next section.

7.3 Examples

Barab and colleagues (e.g., Barab et al., 2010) combined DBR and socially responsible design to create an intervention that would help students develop their identity both as individuals and members of their community along being educated to be knowledgeable citizen of the world. The project spanned over five years and included several iterative components, ultimately designing a video game that became known as Quest Atlantis (QA), with teachers, students, community members and web designers as part of the research partnership. Key to the DBR methodology, the research group developed a theory about transformational play: that video games can serve as effective mediums for deep and sustained learning by providing engagement not possible in the classroom (Barab et al., 2010). This theory has also been used by other researchers (e.g., de Sousa et al., 2018) as a framework to their work. One project that grew from the theory developed by Barab and colleagues is the extension of the *Adventures of Jasper Woodbury* mathematics curriculum work that has been undertaken by Gresalfi and her colleagues (e.g., Gresalfi & Barnes, 2012). Like the development of Quest Atlantis and the development of the original *Adventure of Jasper Woodbury* series (e.g., CTGV 1992; 1994), Gresalfi has undertaken the new work related to the *Adventures of Jasper Woodbury* by engaging in a DBR approach that looks at the relationship of student mathematics learning activities (Type B) and student mathematics learning outcomes (Type A), but also incorporates variables of Types C (interactive mathematics teacher activities) and G (individual student characteristics, abilities, and personal qualities). Her group designed the Boone's Meador mission as an activity that provides insight into Type B variables (student mathematics learning activities). In Boone's Meadow, students are tasked with making calculations and decisions regarding how to reach a destination in order to save an endangered eagle. The game measures student learning outcomes based on student's responses and decisions made throughout the activity, shedding insight into Type A variables (student mathematics learning outcomes). The game includes feedback that is meant to reflect the actions of a teacher, thus including Type C variables (interactive mathematics teacher activities), as well as how that feedback affects the activity of the game (Type B—student learning activities) and student learning (Type A—student learning outcomes). Gresalfi and Barnes (2012) did two iterations of DBR to design and explore the effect of consequential feedback, which is feedback that is embedded in context and requires students to evaluate their mathematical reasoning based on the outcome of their decisions. The two rounds of implementations spanned across two years; data sources included videotapes of discussions and student work. Several rounds of data analysis were done using both a priori and emergent codes. The team saw an increase in the use of mathematical justification and critical engagement when feedback was embedded in context and given prior to the end of the game.

An extension of DBR that arose in the early 2010s is Design-Based Implementation research (DBIR), in which implementation becomes the vital focus of theory development and analysis (Penuel et al., 2011). DBIR often includes the combination of learning sciences research and policy research. One such example of DBIR is the work of Cobb and colleagues (e.g., Cobb et al., 2013), who partnered with four urban schools to improve the quality of mathematics instruction with an 8-year project titled Middle School Mathematics and the Instructional Setting of Teaching (MIST). The focus of improving mathematical instruction was broken down into increasing the learning of conceptual understanding, justifying solutions, and explicit connection between multiple representations. The researchers believed that a reorganization of teacher's instructional practice was necessary for these improvements to occur.

The research partnership consisted of school and district leaders, math coaches, teachers, and researchers. The iterative design process consisted of yearly cycles of data collection, analysis and feedback: they documented district's improvement strategies, collected and analyzed data to assess the implementation of the strategies, and recommended revisions of strategies for the following year. Additionally, a secondary level of focus was on gathering data to test and revise conjectures about supports and accountability measures that the research group had generated from literature. Examples of data collection methods include audio-recorded interviews, district organizational schedules, evaluation forms, online surveys, classroom observations, and student achievement data. Cobb's team (Cobb et al., 2003) analyzed their recommendations for each district, looking patterns and similarities to find potential generality. After the third year of data collection, retrospective analysis was conducted to provide evidence for conjectures about major components of their emerging theory of action for instructional improvement in mathematics.

As part of the theory development, the researchers designed, tested and modified conjectures about instructional improvement, more specifically on methods of both supporting and holding teachers accountable for reorganization of practices. They developed an interpretative framework that captured four general supports that the districts used in the improvement strategies: new positions, learning events, organizational routines and tools. The research on the ways that mathematical instruction improved map to Medley's Type E variables, focusing on teacher competency, knowledge, and skills. The researchers also attended to the ways in which a change in teacher's instructional practice, including variable D, pre and post-active mathematics teacher activities, and variable C, interactive mathematics teacher activities, was tied to mathematics teacher's competence, knowledge, and skills (Type E). Further, the MIST research team was able to provide recommendations for the district on how to support teachers. The four recommendation areas focused on variables of Type I (external context variables) and J (mathematics teacher development and experience). Therefore, MIST was able to study the relationship of variables of Type E (teacher's competence, knowledge and skills), D (pre and post-active mathematics teacher activities), and C (interactive mathematics teacher activities) by surveying and observing teachers. Recommendations are also focused on variables Type I (external context variables) and J (mathematics teacher training and experience) as the research team partnered with the school leaders to influence those categories outside of teacher control. MIST has continued to work with schools as partnerships in implementing strategies to improve math instruction and teacher practices (see Cobb et al., 2018 for more detail).

8 Cultural Historical Activity Theory

8.1 What is It?

Cultural-Historical Activity Theory (CHAT) is a theoretical framework for conducting sociocultural research. CHAT supports the analysis of human interaction while considering how an individual or group of individuals and their interactions with the environment affect their activities (Kuutti, 1996; Cole & Engeström, 1993; Engeström, 1993). The basic idea of CHAT is that humans should not be separated from their participation in various activities; therefore, rather than focusing on the individual as the unit of analysis, CHAT instead focuses on the activity in which people participate (Cole & Engeström, 1993; Engeström, 2004). Thus, the unit of analysis includes both the setting and individuals. The activity system refers to a collective concept—it is object-oriented, designed to think about the phenomenon in terms of the inner relations of activity and collaborative relationship between people (Roth, 2012). For the purposes of this chapter, we are focusing solely on the psychological aspects of CHAT and not on economic or materialist aspects. We assert that is most productive for this chapter's focus.

CHAT was initiated and developed by Russian theorists who saw behaviorism and analytical psychology lacking in its ability to describe cultural realities. The pedigree of CHAT can be traced back to dialectical materialism, and then to Lev Vygotsky who founded the first-generation of activity theory in the 1920s, centering it around his core idea: cultural mediation that is graphically expressed as a triangle with subject, object, and mediating artifact/tool comprising the vertices (Cole, 1998). The basic elements could be described as:

- Subject—The individual or subgroup involved in the activity.
- Object—The problem space or recipient of action to which the activity is directed to be molded or transformed in reaching the outcome that is sought.
- Mediating Artifacts/Tools—Internal mental signs and external physical objects that facilitate and support thinking processes and regulate interaction between the individual and the world. The artifact is "an aspect of the material world that has been modified over the history of its incorporation into goal-directed human action" (Cole, 1996, p.117)

Beyond the prevailing behaviorist theories about the stimulus–response association at that time, Vygotsky's mediation triangle, as a semiotic process between subject, mediating artifact, and the object of an activity, was a revolutionary way individual make meaning of the world (Cole, 1996; Cole & Engeström, 1993; Yamagata-Lynch, 2007).

The first-generation theory was critiqued, because the unit of analysis still focused on individuals. To overcome it, Alexei Leont'ev (1981), Vygotsky's colleague and disciple, along with his colleagues, created a second generation of CHAT, which took into account inter-relationships between the individual and the community, history, context, and interaction of the situation and activity. According to Leont'ev (1974): "activity is…a system possessing structure, inner transformations, conversations, and development" (p. 10). Thus, the consequences of events and activities that occur during the activity can qualitatively change the participants, the participant's participation purpose and motivation, the social environment of the activity, and the activity itself (Rogoff, 2008; Rozin, 2004; Yamagata-Lynch, 2010).

According to Engeström (2004), Leont'ev never graphically expanded Vygotsky's original model to illustrate a collective activity system. In addition, Leont'ev and his colleagues did not adequately address the methodological challenges for capturing, analyzing, and presenting activity-based data. To address these shortcomings,

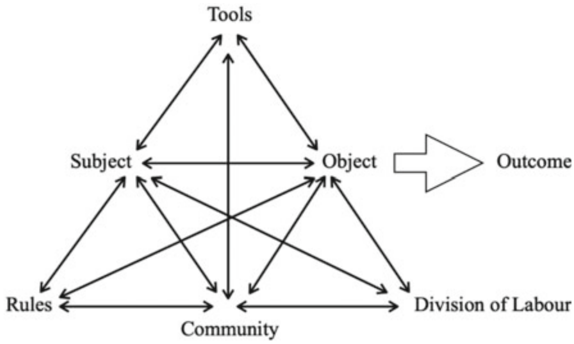

Fig. 1 Expanded mediational triangle (EMT) (Engeström & Cole, 1993)

Engeström created the third generation of CHAT, offering a foundation for understanding and designing learning as a transformation of human activities and organizations. Engeström and his colleagues developed CHAT as an analytical framework by introducing a descriptive model of activity, which can be used in analyses of complex qualitative data. Compared to other sociocultural learning theories, Engeström's theory of expansive learning puts the primacy "on communities as learners, on transformation and creation of culture, on horizontal movement and hybridization, and on the formation of theoretical concepts" (Engeström, 2010, p.2). Cole and Engeström (1993) further detailed the representation of modeling human activity as a system form in the diagram of expanded mediational triangle, shown as Fig. 1. This is the triangle that typifies CHAT research.

The triangle provides an organizer to support researchers in mapping complex human interactions that take place in collective settings. The uppermost sub-triangle is identical to Vygotsky's basic structure of mediated action. In addition to the basic components of Subject, Tools and Object presented in the basic first-generation triangle, the expanded mediational model also includes the following three elements:

- Rules-norms, regulations, convention and gnorms, regulations, convention and guidelines that afford or constrain action and interaction within an activity system.
- Community-multiple individuals and subgroups involved in an activity.
- Division of Labor-distribution of work and responsibilities between members of the community.

The Rules, Community, and Division of Labor in the bottom portion of the triangle model add the sociohistorical collective nature of mediation that was not addressed by Vygotsky (Engeström, 1999a, 1999b). The outcome is the results or consequences that the subject finds once the activity is completed (Engeström, 1993, 1999a). Engeström (1999a) explained that the relationship between components of an activity system is two-way as people not only use instruments, but also renew them, they not only use rules, but also reformulate them.

The interactions among the components of the triangle model highlight tensions that are inherent in human activities; researchers find tensions in activity systems when elements from one or more components pull participants away from achieving the purpose of the activity thus cause changes in activities, so tensions may either promote or hinder human activities. (Cole & Engeström, 1993; Engeström, 1993, 2004; Yamagata-Lynch, 2003). We argue that from the perspective of Medley's (1987) variables, CHAT is appropriate for looking at relationships between any subset of the variables, depending on the data to be collected. Because CHAT was developed to consider complex systems, it is particularly suited to the task Medley was conceptualizing in the development of the presage-product-process perspective (Medley, 1987).

8.2 Benefits and Limitations

The primary benefit of CHAT is its inherent ability to make sense of a complex system in a way that accounts for the actors and mediators at work in that system (Yamagata-Lynch, 2010). As exemplified in the examples below, CHAT provides a means for making sense of external context variables and the effects they have on instructional activities for teachers and students. This is important if the field wants to extend beyond Mendeley's (1987) original assertion that only two adjacent levels of variables can be considered at one time.

The limitations of CHAT are important considerations. First, it is not appropriate for considering human thinking, as it relies on observable activity (Yamagata-Lynch, 2010). This has implications for the kinds of growth that can be considered, how knowledge is characterized, and other elements of consideration that can only be observed by proxy. Further, the triangle model, while supporting sense-making about human activity systems, also oversimplifies those systems (Yamagata-Lynch, 2003, 2010). That is, complex human interactions are summarized to the point that they are "…not as rich and complex as real experiences" (Yamagata-Lynch, 2010, p. 33). Finally, CHAT is complicated to learn. This is because it requires the researcher to be proficient in qualitative methods, to understand and honor the complexity of collecting trustworthy data, and the ability to bring all of that together within a very specific theory (Yamagata-Lynch, 2010).

8.3 Examples

One of the challenges of STEM education is to integrate activities, content, and tools in a meaningful in-class activity. Using CHAT models as a basis for analyzing learning, teaching, and in-class interactions between different subjects calls for the transformation of authentic scientific/mathematics practices into classroom activity systems. Here we provide two examples that have relied on Engeström's theorizing of CHAT. We also invite the reader to look at the work of Schmittau, who used cultural-historical theory as put forward by Vygotsky to make sense of student's mathematical learning (e.g., Schmittau, 2004, 2005, 2011).

CHAT has become an important lens in mathematics education research because it has “power to deal with complexity in educational systems” (Jaworski & Potari 2009, p. 222). To build on early research with use of activity theory in mathematics teaching–learning, as well as with a focus on the classroom tasks and their related macro-social setting, Jaworski and Potari (2009) considered teaching as activity in their study in the 10th grade classroom of a UK secondary school where students in this grade group are considered “lower achievers”. They used CHAT to consider the role of the social framework within which classroom teaching is situated. They had two primary goals. The first was to understand the relationship between teacher-student interactions and the ways in which cognition is evident in classroom dialogue. The second was to analyze the relationships between classroom interaction and cognition within the broader cultural context in which learning occurs. They employed triangles from EMT to characterize the “subject” to be any teacher or pupil learning in this setting, each with their goal or object for their activity.

Specifically, Jaworski and Potari (2009) analyzed teacher-student interactions through classroom dialogue, which they viewed as a micro-analysis. In an episode offered by the authors, the teacher, Sam, had planned a didactical inquiry including in-class activities and relevant resources. In the implementation of this plan, Sam met with some “tensions” (p. 229). For example, students who had not done their homework completely derailed Sam’s lesson plan. Jaworski and Potari suggested explanations behind the homework issue. They talked about the task that teacher assigned to students: from a teacher’s perspective, the activity seemed “logico-mathematical” and reasonable in “didactical communities”; however, for student peer and family communities, it is “strange and unreasonable”.

The representation of the application of CHAT to allow analysis from both teacher’s and student’s perspective. For example, through their analysis, Jaworski and Potari (2009) determined: teacher’s object could be “understanding of basic statistical terms and associated concepts,” while pupil’s object would be “classroom survival”; teacher’s rules could be “teacher/student authority structures,” while pupil’s rules are “homework expectations within the school”; and the outcomes for teachers are “Non achievement of object due to pupils not taking the required responsibility, tension in the classroom”, for pupils are “survival by ignoring terms of homework, contravening rules and contributing to classroom tension” (p. 231). By illustrating the descriptive power of CHAT for making sense of the observable activities in mathematics classrooms, the researchers framed teacher’s mathematics teaching as inconsistent with their socio-cultural histories. Further, they found that the teaching did not match non-dominant student’s learning. This highlighted, for the researchers, the lack of opportunities for mathematics teachers to challenge privilege-oriented activities. Without these opportunities, many well-intentioned mathematics teachers may unconsciously continue to perceive, explain, and respond to the classroom activities and specific learners through the dominant discourse system, which triggers the equality that they originally desired to abolish. (Jaworski & Potari, 2009).

When we place this study in Manizade et al.’s (2023) adaptation of Medley’s (1987) framework, we can see the role of external context variables (Type I), where the teacher’s preparation of instructional materials (e.g., homework designed for students), school rules about homework, and student's parent's supportive attitudes toward homework are all factors that contribute to student's responses to homework. In Sam's example above, his preparation for class belongs to Type D (mathematics teacher's competencies, knowledge, and skills), including his requirement for students to do pre-work before class, and his design of in-class activities based on student's pre-work. The relationship of variables of Type I (external context variables) to Type D (pre- and post-active teacher activities) show up as tensions between *tool* and *community* in EMT framework. Then, in the analysis of the student's and teacher's perspectives, we can see interactions between variables of Type C (interactive teacher activities) and Type B (student learning activities), but unlike other research in which the influence is only considered in one direction (e.g., from Type C to Type B), CHAT allowed the researchers to understand the relationship in both directions—that is, the student's perspectives on the learning activities and the teacher's perspectives on the student's characteristics, abilities, and personal qualities (Type G) through the lens of their interaction with the learning activities. In summary, this example of mathematics teaching–learning interaction in the CHAT framework shows the power of this new method for addressing the interactions between and among the presage-process–product variables in Medley's framework.

In a separate study, Black et al. (2010) offered new insights into student's identity development by exploring an implicit mediation: they drew on Leont'ev's approach to gain an understanding of "self" related to mathematics. Mediation has a complex and abstract nature, studying an unintentional and less obvious object, like identity development or mental functioning, could be implicit. Driven by the interest in student's perception of themselves in relation to future aspirations, particularly their mathematical identity shifts, Black et al. (2010) conducted post-observation interviews with Mary and Lee (aged 16–17 years), two students studying advanced-subsidiary level (AS level) mathematics in England, to explore the relationship between learner's identity and mathematics. Black and colleagues (Black et al., 2010) adopted the methodological tool "leading activity" adopted from Leont'ev, which framed their understanding that activities become leading and can trigger a new activity when new motives are generated that surpass the original motive. In this work, the researchers found that satisfying mathematics-learning experiences implicitly mediated a "leading identity" that affected student's career choice, for example, in Mary's case, her identity also represented as her motive for studying mathematics is 'vocational (get a good job)'; however, in Lee's case, his focus on study as an activity is mediated by both his identity shifting and motive for attending a university. As such, Black et al. (2010) built on CHAT theories by presenting a relationship between self-identity and one's motive to engage in activity.

Considering the Black et al. paper (2010) from the perspective of Medley's framework, we can see the interaction of variables of Type G (individual student characteristics, abilities, and personal qualities) with Type A (student learning outcomes). In Mary's case, her engineering project experience as a leading activity significantly drove her to her future potentiality. Meanwhile, her self-awareness of the needs as "I like hands-on stuff" with some other positive aspects in her personality contribute to her motivation to become an engineer. In contrast, Lee did not value mathematics as much as Mary did, so his purpose for studying mathematics and engagement with the subject was less meaningful than Mary's (Black et al., 2010). This analysis highlights how Type G variables (individual student characteristics, abilities, and personal qualities) may impact learning outcomes.

9 Quantitative Ethnography

9.1 What is It?

Quantitative ethnography (QE) is an emerging approach to research that attempts to bring quantitative and qualitative analysis of data into the same conceptual framework (Shaffer, 2018a, 2018b). That is, QE draws from the tools, perspectives, and approaches of both qualitative traditions and quantitative traditions to create a mixed methods approach that is philosophically consistent with both. This is a research approach that builds from the emergence of Big Data, which has allowed the collection of data that can be simultaneously as rich as traditional ethnographic data while being collected in quantities previously reserved for only the largest studies (Shaffer, 2017, 2018a; Wooldridge et al., 2018). While many approaches to working with big data have focused on statistical analysis, QE offers a different approach.

QE has been developed grounded in the belief that learning is an interpersonal activity. Learning is conceived of as making meaning of the world in a way that is consistent with how a particular group makes meaning of the world (Shaffer, 2018a). That is, learning is about induction into a community of practice (e.g., Lave & Wenger, 1991; Wenger, 1998) and, thus, it is about learning the Discourse of that community or culture. Discourse, in this case, refers to Gee's (2014) notions of "Big D" Discourse, which is any culture's way of being, including people's ways of talking, listening, interacting, believing, valuing, and feeling. In this case "small d" discourse becomes the observable behaviors through which researchers can gain insight into Discourse, as Discourse cannot be readily observed. "Small d" discourse is what people actually say and do. Thus, when we use QE to assess and understand learning, we are looking for the ways in which participants express their changes as they are inducted into a community of practice. This focus on induction as the outcome of learning makes QE particularly appropriate for considering student learning outcomes (Type A) in light of the instructional environment variables including student activities, interactive teacher activities, and pre and post-active teacher activities (Types B, C, and D) while considering many contextual variables, including Type H (internal context variables), Type I (external context variables) and Type J (mathematics teacher development and experience).

In quantitative ethnography, the research process is distinctly and necessarily mixed methods (Wooldridge et al., 2018). The data collected are rich in nature, just as they would be in traditional ethnography. They are collected using rigorous qualitative methods and may include traditional qualitative data such as observations and interviews, or newer forms of data collection such as data collected by the computer as students work together in a virtual environment. Such data could include clickstream data as well as full transcripts of interactions. The analysis of the data is where we start to see the mixed methods nature of the approaches. For example, in epistemic network analysis (ENA; Shaffer et al., 2009, 2016; Shaffer & Ruiz, 2017; Shaffer, 2018b), the data are coded using frameworks from discourse analysis (Gee, 2014), which structures analysis by breaking data into segments that are typically a single utterance and joining those segments into logical chunks called stanzas. Then, ENA draws from traditional qualitative research, particularly grounded theory (e.g., Charmaz, 2014) to code data using approaches such as those used in grounded theory or inductive analysis (e.g., Maxwell, 2013) to create a coding scheme which is then applied to every segment. Once this is done, ENA draws from social network analysis (e.g., Robins, 2015) to mathematically create a visual display of the interactions between codes. The visual display (e.g., Fig. 2), shows the prevalence of single code (represented by a node) through the size of the node, and it shows the strength of the connections between nodes through line thickness. In this way, the visual shows those ideas (codes) that co-occurred in a single statement, which is a proxy measure for the codes having some kind of connection to each other for the person speaking. From this visual, additional statistical analysis, such as t-tests to determine whether particular groups are significantly different, or additional qualitative analysis, such as looking at all of the instances in the transcripts captured by particular node connections can be pursued.

As an example, we present two ENA maps from the first author's dataset in Fig. 2. This data was collected as part of a larger study focused on how middle school mathematics teachers understand proportional reasoning. The two teachers featured here (Autumn and Patricia—all names are pseudonyms) are representatives of the larger pool of 32 teachers. Each teacher responded to a number of items related to

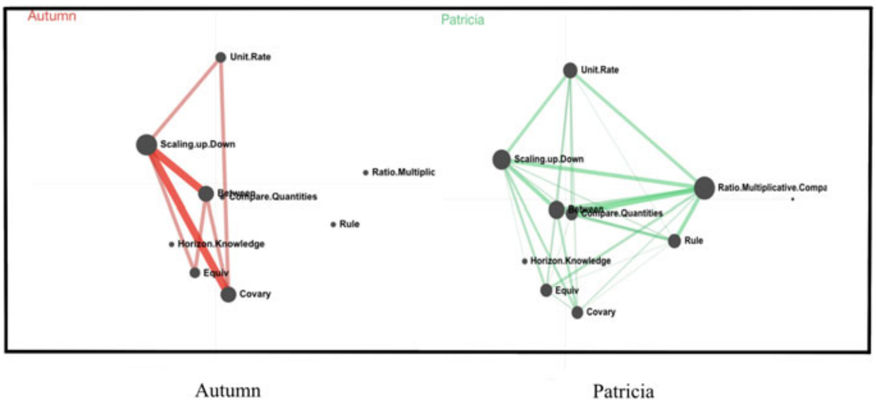

Fig. 2 Two teacher's ENA plots showing the knowledge resources they used for reasoning about proportional situations

proportions that were designed to help us understand how they reason about proportional situations. Part of the data was collected using a face-to-face clinical interview (Ginsburg, 1997) and part was collected using a think-aloud protocol mailed to the participants for which they used a Livescribe pen that captured their talking and their writing to create a record of their thinking. The coding scheme was developed using a grounded theory approach (Weiland et al., 2020). The figure shows the ENA mapping of our analysis of Autumn and Patricia's responses to the items. In these maps, each node shows a particular mathematical understanding that was included in their response, and the lines connecting the nodes show where they discussed those mathematical ideas in the same utterance. For example, in Autumn's map, we can see that Comparing Quantities, Scaling Up and Down, and Covary were the most commonly used knowledge resources because those nodes are largest. Further, we can see that she often talked about Covary and Scaling together and she talked about Scaling and Comparing Quantities together frequently. In contrast, we can see that Patricia relied more on Ratio as a Multiplicative Comparison and Scaling Up and Down. However, she did not demonstrate strong connections between the knowledge resources that were as frequent as Autumn demonstrated. By using ENA, we can see different patterns among teacher's data, which helps us understand what knowledge they access while solving problems and where there may be opportunities for professional learning. While this may appear to be only focused on variable Type E (mathematics teacher's competencies, knowledge, and skills), we would argue that it is also capitalizing on Type C (interactive mathematics teacher activities) and Type D (pre and post-active mathematics teacher activities) because we have found that situating conversations of teacher knowledge in the context of the decisions teachers make about students and instruction provides additional insights into the teacher's knowledge of the content as it relates to their teaching. Thus, doing this kind of research relies on the interaction between Types C, D, and E to understand teacher knowledge in context.

9.2 Benefits and Limitations

Quantitative ethnography is unique in its approach to using large amounts of data to create thick, rich accounts of the situation. From the perspective of Medley's (1987) framework, this method allows us to look explicitly at the interactions within a single element or to look at interactions across elements depending on the framing of the research questions. In fact, using epistemic frame theory (Shaffer, 2004, 2006, 2009, 2018b), which is one set of axioms that can form the basis for the framing of the study and the analysis of data, one would specifically consider how the elements of learning or teaching are situated within the culture of the classroom over time. Learning would only be conceived of as interpersonal, meaning that the interactions between teachers and students would be one site in which one would look for changes in the nature of discourse. QE allows the collection of large amounts of rich data that can be analyzed in ways that capitalize on both statistics and qualitative approaches.

Despite its origins as an assessment tool for learning in simulation environments, QE has evolved to be useful for other kinds of analyses, such as the analysis of teacher knowledge shown above. Thus, while it is grounded in a clear theory of learning, the methods can be used in other ways. This is consistent with other methods as well, including grounded theory.

Depending on the research question and focus for coding, there are tools that can help with the initial coding of data for QE. If the data being analyzed can be coded by the computer, that can save considerable time as coding a large body of data can be very slow. Studies that adhere more closely to the ideas of measuring discourse as a means for understanding Discourse, for example, could include coding of keywords and concepts that could be captured through computerized coding. In contrast, work like that done by the first author cannot benefit from computerized coding, because interviewees do not necessarily use consistent language to express certain understandings and because some keywords are used in a variety of ways ranging from ways that indicate strong understanding to ways that do not. Thus, for some research, QE can be very time intensive, while for other research it is less so.

9.3 Examples

Much of the initial work with ENA that has led to the development of QE was focused on learning games designed to help learners assimilate into the community of practice relevant to the game. For example, Nephrotex (e.g., Arastopoor et al., 2012) and RescuShell are two simulations that provide engineering students with virtual internships during their first year of an engineering program. In each, students are presented with a design problem that they work to solve. In one recent study of these environments (Chester et al., 2015), the researchers wanted to know what students learn from a course based entirely on working in these two simulations. They collected data from 50 students across the semester. Data collected included pre and post-surveys built into each simulation as well as all of the student's chats, emails, notebook entries, and work products entered into the systems throughout the semester. Data were analyzed using an engineering epistemic frame that had codes in the categories of knowledge, skills, identity, values, and epistemology. The researchers were able to analyze these data using ENA to measure student's development within the engineering epistemic framework. From the analysis, they learned that participating in two virtual internships was more effective than participating in just one. While participation in one simulation led to connection making between skills and knowledge, participation in a second simulation led to additional connections with knowledge of the client and epistemic aspects of engineering, which the authors assert are important aspects of thinking like an engineer. They also found that students were more satisfied with the course at the end of the second simulation than at the end of the first, though student satisfaction was predominantly positive for both. From the perspective of the framework of research on teaching mathematics adapted from Medley (Manizade et al., 2023), this is research focused on Type A variable (student mathematics learning outcomes). But, it uses Type B (student learning activities) and Type E (teacher's competencies, knowledge, and skills) variables to explore the student learning. Specifically, the researchers used student's evidence from their activities (Type B) to determine the learning outcomes (Type A); and the measure of those learning outcomes was based on how similar the student's connection-making had become to the instructor's (Type E). Because moving students to think in ways that are consistent with the instructor is the explicit goal of these simulations, the planning (Type D—pre- and post-active mathematics teacher activities) and interactive mathematics teacher activities (Type C) are developed as explicit stepping stones connecting teacher knowledge to student knowledge.

While popular in the learning sciences, QE is only beginning to emerge in mathematics education. One example of a mathematics education implementation of QE is from pilot work completed by the first author and her colleagues (e.g., Burke et al., 2012; Orrill & Shaffer, 2012). That research focused on the knowledge resources in-service middle school mathematics teachers exhibited as they reasoned about a number of mathematics tasks. In this work, Knowledge in Pieces (e.g., diSessa, 2018) was used as a conceptual framework to drive the identification of fine-grained understandings being used by the teachers. The focus on this work was determining whether there are differences among the relative connectedness of the knowledge resources for the teachers. The hypothesis being that teachers who exhibit more connections between and among their knowledge resources may be better situated to engage with a wider range of student ideas. The work showed that there were unique patterns of knowledge resource used among the teachers and suggested that areas worthy of further research included consideration of teacher's classroom experience (e.g., the development of pedagogical content knowledge) and the relative strength of teacher's mathematics knowledge. As noted above, this line of research embeds teacher's competencies, knowledge, and skills (Type E) in the work that teachers do, which is interactive, pre-active, and post-active mathematics teacher activity (Type D and Type C) to understand how it impacts student's opportunities to learn.

10 Technology for Research

As we alluded to in the discussion of quantitative methods, technology has revolutionized aspects of the research enterprise. It has changed the kinds of data we can collect, which changes the kinds of questions we can ask. Suddenly, we can access new data through tracking devices (e.g., Lee et al., 2015), uncover thinking in news ways by collecting data using different tools (e.g., Hickman, 2015), and engage in mathematical thinking in different ways as technology allows us to interact in more tangible ways this those ideas (e.g., Hegedus & Roschelle, 2013). While a comprehensive review of the ways in which technology has shaped presage-process–product research is beyond the scope of this chapter, we offer three examples of the ways in which technology has fundamentally shaped the research that can be done. We first look at eye tracking, which allows the capture of data previously unavailable to researchers, thus being appropriate for questions about the interactions between student learning outcomes and learning activities (Types A and B) with the interactive mathematics teacher activities (Type C). Then, we discuss the use of dynamic geometry software as one tool that is useful for better understanding how people reason about geometric situations as it allows the researcher and participant to move away from discussing a single example, to instead potentially focusing on an entire class of examples. This allows us to consider the interplay of variable Types A, B, C, and D, but even more, it allows us to ask different questions about student learning outcomes (Type A) and teacher's competency, knowledge, and skills (Type E) than we can ask without dynamic environments. We end with a discussion of 360° video, which opens opportunities for both teaching and researching teaching, and has supported researchers in adding in important ways to the literature on teacher noticing. As with CHAT, 360° video opens an array of possibilities for the researcher to examine all the variables acting together to create the learning environment.

11 Eye Tracking

11.1 What is It?

Eye-tracking technology has made it possible to track and record the eye movement of people looking at screens or paper, which provides data focused on what the person is attending to on the screen. Eye tracking was initially used primarily in reading research but has been gaining popularity in the field of mathematics, particularly being used to analyze multimedia learning processes. Multimedia learning can be referred to as creating mental models from resources that contain both verbal, both spoken and written, and pictorial representations, such as graphs, animations, or tables (Mayer, 2005). Eye trackers can either be attached to a computer monitor or to a head mount wore by the participant. In a recent review of 161 eye tracking studies in mathematics education research (Strohmaier et al., 2020), almost all of the studies used a computer monitor attachment. The data provided from eye trackers is usually in the form of coordinates, which are then categorized into groups of events using automated or manual algorithms (Strohmaier et al., 2020). The information gained from eye tracking can be used, for example, to improve the design of instructional material or answer questions about the differences between novice and experts. Eye tracking provides insight into the relationship between variables of Type B (student learning activities) and Type A (student learning outcomes). More specifically, it allows researchers to better understand which aspects of the screen (instructional activity) students focus on as they complete their work. Assertions can be made about the design of the activity and how it influences student learning. Implicitly, researchers can study individual student characteristics, abilities, and personal qualities (Type G) and internal context variables (Type H) such as patterns in where student attention is given.

11.2 Benefits and Limitations

Eye tracking has allowed researchers to gain insight into visual attention, which is often done too quickly and even subconsciously for participants to register and report on; researchers now have access to data that is not observable to people. This technology provides objective and numerical data that can be used both in qualitative and quantitative research; this unique information on what is being attended to, for how long, and in what order can be used in numerous ways to answer questions that were unable to be addresses previously (van Gog & Scheiter, 2010).

With any use of technology comes limitations. There can be data loss, particularly in very young or old participants; about 10% of data can be blinks and saccades, which provide no valuable information. Additionally, accuracy of the eye tracking device can be hindered with head-mounted devices (Strohmaier et al., 2020). It is noteworthy for researchers to be aware that eye tracking only reports data on what the participant is attending to; there is no data on that can give any explanation as to *why* the participant is looking at certain places. Therefore, researchers must rely on making inferences about any cognitive processes underlying the movement.

11.3 Examples

Eye tracking can be used to provide data for numerous research purposes. For example, it can be used to study how participants split their attention when presented with texts and diagrams. For example, Andra et al. (2015) investigated difference between how students look at formulas and graphs of linear equations, thus linking variables of Type A to Type C, with implicit attention to Type B. The review mentioned previously (Strohmaier et al., 2020) found that a majority of the mathematics studies covered the topic of numbers and arithmetic, studying, for example, how participants represent and process numbers, calculations, and equations. The topic of geometry, particularly shapes and form, was another common topic for researchers to investigate (See Strohmaier et al., 2020 and the special issue of *Learning and Instruction* (2010, 20(2) for mathematics examples).

12 Dynamic Geometry Software

12.1 What is It?

To enhance twenty-first century student's learning process and academic performance, a pioneering technology development, dynamic geometry software (DGS) has become a main feature that acknowledges the idea of 'interpretative flexibility' (Ruthven, 2018). By borrowing the idea of creating dynamic rather than static graphics from contemporary drawing software, it is possible to drag the objects such as points, line segments, or circles of the graphics while retaining the defined properties, and the dynamic feature can be reflected in some "transformation" manipulation, such as translations, reflections, rotations, and dilations, with the help of mouse or tracker-ball on a laptop. DGS has been regularly used worldwide for teaching and learning geometry, with software like GeoGebra, Geometer's SketchPad, and Cabri Géomètre being common in many mathematics classrooms. More and more, it is being used to uncover understandings about mathematics concepts in ways that attend to transformations, thus allowing the researcher and participant to have something visual to discuss as they consider the mathematical ideas. DGS can be used to better understand student learning outcomes (Type A) as well as the interaction between Type B (student mathematics learning activities) and Type A variables. Research on teacher knowledge, such as that discussed below, can also focus on Type E (mathematics teacher's competencies, knowledge, and skills) and, if a researcher wanted to understand the ways in which DGS can be used to promote learning, a design that connects variable Types E (teacher competency, knowledge and skills), D (pre and post-active teacher activities), C (interactive teacher activities), B (student learning activities), and A (student learning outcomes) could be developed.

12.2 Benefits and Limitations

In Geometry class, DGS could support children's learning transition from "because it looks correct" or "because it works in these situations" to robust mathematical understanding of the geometric situation (Jones, 2000). To be specific, applying DGS can provide opportunities for students to find patterns in abstract geometrical graphics, so they can conceptualize mathematical ideas, such as invariance, or perceive mathematics rules, such as the relationship of the leg lengths in triangles, with less vagueness. For example, purposive manipulation like dragging along a circle can help make a defined property—the unchanging measure of an angle of circumference—comprehensible and convincing to students. Despite the benefit of visualization and facilitation, teachers hold different perspectives on the efficiency issue of students using software in class. In some situations, only teachers use software for in-class presenting because they concern that students would invest in-class time to get familiar with software operation (Ruthven, 2018).

As a research tool, DGS opens new ways to engage teachers and students in explaining how they understand geometric concepts. Rather than being limited to *describing* a phenomenon based on a drawing on paper, the interviewer and interviewee can engage in *showing* each other what they mean. This opens a pathway for richer understanding of participant's knowledge.

12.3 Examples

Martinovic and Manizade (2020) explored teacher's thinking through the use of DGS by examination of both the teacher's written work and GeoGebra sketches from 23 in-service secondary school teachers in the USA. How these teachers visualized and verified the trapezoid area formula conjectures in GeoGebra is quantitatively as well as qualitatively analyzed as "empirical proofs" (p. 3) of their strategies and connection to teaching. From the qualitative analysis, the authors identified four distinct strategies: eyeballing, measurement, constructions, and written statements, and they found that the teachers used a combination of these four strategies. They also found out teacher's "misconceptions" (p. 16) that were magnified in the process of using technology, and some operation failure of that some teachers may treat DGS as a paint software. Overall, this study contributes on the teacher's strategies of visualizing and verifying the trapezoid area formula conjectures, also widen the scope of potential research on teacher's knowledge in the context of geometry class. This study focused on Type E (teacher competencies, knowledge, and skills), with implications for improving and modifying Type J (mathematics teacher development and experiences).

Nagar (2019 ; Nagar et al., 2022) found four categories of invariance that teachers were able to identify by engaging them with a series of four DGE protocols. He found that teachers did not discuss invariance at all without prompting, but when prompted, they were able to use the DGE to highlight important aspects of the geometry. This was particularly interesting given the notoriously difficult task of uncovering invariance in other work (e.g., Laborde, 2005). Consistent with Nagar's speculations that this work will inform how we teach students to better understand geometry, we would consider it research focused on variable Type E (mathematics teacher competencies, knowledge, and skills) with implications for variable Types D (pre- and post-active teacher activities), C (interactive teacher activities), and B (student learning activities).

12.4 360° Video and Other Full-Room Video Capture

12.4.1 What is it?

The emergence of affordable video tools and better computer programs for controlling video has opened opportunities for research classrooms to be built that are designed for researchers to capture the entire experience of the classroom. These rooms can include multiple video cameras and multiple microphones. Sometimes, they include a control room from which a researcher can control the data collection. The goal of these rooms is to capture as much data as possible in real time.

At the same time as these teaching and research labs are emerging, new technologies are making it possible to collect full-room video in other ways, too. For example, some researchers are using video cameras that are designed to capture 360° views rather than just the normal framing of video. Other researchers are using tools that allow you to control a tablet computer to follow a speaker and record that person as they teach or present. This can allow remote data collection as well as collection of data using a number of devices in a single setting. Similarly, some researchers have used wearable video cameras, such as GoPros, to see what each participant in a study can see (e.g., Sherin et al., 2008). Like CHAT, this kind of research is rich in the research opportunities it opens. We would argue that any variables, except Type F (pre-existing mathematics teacher characteristics) and Type I (external context variables) could be studied using this technology depending on the design of the study.

12.5 Benefits and Limitations

While the benefits of capturing classroom activity this way are myriad, it is not true that the data is unbiased. As with any data, there is always bias in video data because a human has made a set of decisions based on a set of criteria for data being collected in the space (e.g., Hall, 2000). However, these whole-room approaches to video capture allow something closer to unbiased capture of the experience to happen. While dedicated video suites remain relatively rare because they require dedicated space, the other options (e.g., 360° cameras, GoPros, etc.) are relatively inexpensive and easy to set up in a variety of settings. Clearly, capturing the volume of data made available through this application requires careful attention to research design to ensure that the studies resulting from high volumes of data are doable.

12.6 Examples

In one line of research, the 360° video technology is being used to create the research stimuli. Preservice teachers are asked to watch 360° videos of children learning mathematics as part of lessons on teacher noticing (e.g., Kosko et al., 2020; Zolfaghari et al., 2020). The research has focused both on what the preservice teachers notice in the 360° video versus traditional video views as well as how to promote preservice teacher's noticing of student's strategies. Findings in the Kosko et al. (2020) showed that preservice teachers who used the 360° videos were more successful in noticing both reform-oriented and content-specific aspects of the instruction than those who relied on traditional video views. This research considers the connection between Type B (student learning activities) and Type E (teacher's competencies, knowledge, and skills) variables by looking at them through the teacher interactive, pre-active, and post-active activities (Types D and C) that were implemented.

In another line of research related to video technologies is research focused on determining the most effective uses of the technology for a variety of teaching, learning, and research purposes. For example, van der Kleij and colleagues (2019) undertook a study in Australia to explore the feasibility of using GoPro cameras with iPads to capture student–teacher interactions. Their findings showed that the two technologies used together can be useful and that teachers are able to engage in teacher noticing activities while viewing the videos created with these devices. This was similar to the findings of Sherin et al. a decade earlier (2008), though they only considered wearable cameras and not the addition of the iPad tablets. As with the research above, this study is considering students learning activities (Type B) by watching the planned instruction (Type D—pre-active teacher activities) as it is enacted (Type C—interactive teacher activities).

13 Presage-Process–Product Research in the 21st Century

While this chapter cannot possibly provide an exhaustive discussion of the ways in which research has evolved since the Medley (1987) framework was introduced, we have attempted to offer insights into changes that have shaped the ways in which we think about research, learning, and teaching as well to provide some examples of approaches to research that simply were not available in 1987. Despite the development in methods and tools, it still holds that one cannot study the interactions of the variables without considering the mediating factors (which, often, are other variables from Medley's framework). His assertion that we need to have a clear conception of good teaching, valid instruments, and appropriate data is still at the heart of good research. Perhaps more than in 1987, modern researchers recognize that teaching is multifaceted and there is no single definition of "good teaching"; thus the onus is on the researcher to define the construct and clearly convey the purpose of the research (e.g., Orrill & Cohen, 2016).

Looking back, we can see the emergence of new methods and theories that attend far more to contextual variables and rich details than those commonly in use in 1987. Rather than trying to find particular variables that explain learning or teaching, the field is now more concerned with context and complexity. The research methods and theories that are in use and emerging now reflect that shift. In part, technology is to be thanked for this change as it has made data collection and analysis much easier than it was in 1987.

The advantage of our current research landscape lies in the ways we can challenge Medley's (1987) assertion that "research designed to correlate nonadjacent points is not worth doing" (p. 111). With the tools and approaches we now have available, current researchers have opportunities to think about the relationships between Medley's variables in ways that are more robustly interconnected and less hierarchical. For example, QE and CHAT are both explicitly focused on finding the connections between and among elements within and between the variables. CHAT, particularly, is interested in how teaching, learning, and instruction interact with each other. Similarly, DBR is expressly focused on including the context of the research as part of the consideration of what happened.

At the same time, we challenge Medley's assertion about the necessity of looking only at adjacent variables, as we can now examine the relationships between and among variables in ways Medley could not have imagined. QE, for example, with its ability to draw on large datasets to yield thick, rich description can help us understand connections between the adjacent variables. Instead of being limited to understanding whether Type B student learning activities (Manizade et al., 2023) shape student's Type A learning outcomes, we can now look across large groups of students to find out how those activities shaped learning, how particular groups of students (Type G—individual student characteristics, abilities, and personal qualities and Type H—interactive context variables) interacted with those activities and what was learned, and the influence on individual and group learning outcomes. Teaching experiments offer one model for diving deeply into student outcomes and their relationship to student learning activities by focusing explicitly on student characteristics.

As with QE, DBR also allows us to conduct research on multiple pairs of adjacent variables simultaneously. The unique characteristic of partnering with a variable of professionals allows for each group to provide insight and perspective that can be used to study multiple variables. By partnering with teachers, each DBR project indirectly provides informal development experiences (Type J) that can influence teacher competency, knowledge, and skills (Type E). The initial phase of DBR projects includes exploring and understanding a project in natural context (McKenney & Reeves, 2013), which gives researchers the opportunity to examine internal context variables (Type H) through interviews and questionnaires to better design student mathematics learning activities (Type B).

The examples of DBR studies described in the previous section were able to study multiple pairs of variables. The goal of MIST was to improve interactive mathematics teacher activities (Type C). The research group studied external context variables of the support systems (Type I—external context variables) and teacher pre-active and post-active activities (Type D). They made recommendations for changes in the support system to influence these practices (Type C). Additionally, they observed interactive teacher activities (Type C) to understand how the changes in support impacted what happened in the classroom (student mathematics learning activities, Type B, and student learning outcomes, Type A). They spent several years on investigating variables of individual student characteristics (Type G) and internal context (Type H) variables. The knowledge of these variables directly affected the design of QA, a student learning activity (Type B). Then, the group collected data on the relationship between this activity and learning outcomes.

As shown throughout this chapter, there are explicit and implicit connections between the variables of interest to any research effort focused on the relationship between teaching and learning. When these connections are not explicitly attended to in the research design, the result is research that yields inconclusive or confounded findings. For example, large scale studies of professional development, in an effort to yield clear relationships between teacher development and experience (Type J) and student learning outcomes (Type A) rely on data of Type E (teacher competence, knowledge, and skills) and Type A (student learning outcomes) only without consideration of the steps in between that mediate the effectiveness of the PD (e.g., Wayne et al., 2008, 2011). It is for this reason that conclusive findings from PD are often elusive (Yoon et al., 2007) or very broad, such as those offered by Garet et al. (2011). Attending to only measures of teacher knowledge and student knowledge can also lead to the appearance that professional development had no significant impact on student learning, when the actual relationship is more complicated than those data would suggest (e.g., Garet et al., 2011). This lack of attention to the "in-between" variables is understandable given the immense complexity of understanding not only whether PD impacted student learning (cf., Banilower et al., 2006, 2007). However, exploring the relationships in-between Type E (teacher's mathematics competencies knowledge and skills) and Type A (student learning outcomes) is critical for understanding how, when, and under what conditions teacher professional development can lead to better student learning. The kinds of methods and technologies discussed in this chapter open opportunities for thinking about these connections in new ways.

Our parting observation is that we believe the presage-process–product framework remains a relevant way to conceptualize research. In this chapter, we have attempted to highlight the ways in which the research field has changed over the three decades since Medley offered his framework. We assert that the evolution of research methodologies, research methods, and available technologies has fundamentally changed the landscape in ways that allow the inclusion of multiple variables, rather than limiting them only to adjacent relationships and has allowed more careful consideration of connections and relationships between and among the variables than was possible with quantitative methods and classic test theory.

Acknowledgements Aspects of the work reported here was supported by the National Science Foundation under grants DRL-1054170 and DRL 1813760. 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.

References

Andra, C., Lindstrom, P., Arzarello, F., Holmqvist, K., Robutti, O., & Sabena, C. (2015). Reading mathematics representations: An eye-tracking study. *International Journal of Science and Mathematics Education, 13*, 237–259.

Arastoopour, G., Chesler, N. C., D'Angelo, C. M., Opgenorth, J. W., Reardan, C. B., Haggerty, N. P., Lepak, C., & Shaffer, D. W. (2012). *Nephrotex: Measuring first year student's ways of professional thinking in a virtual internship* [Paper presentation]. American Society for Engineering Education Annual Conference, San Antonio, TX. Paper No. AC 2012–3103.

Baker, F. R. (2001). *The basics of item response theory* (2nd ed). ERIC Clearinghouse on Assessment and Evaluation. https://files.eric.ed.gov/fulltext/ED458219.pdf

Baker, F. B., & Kim, S. H. (2004). *Item response theory: Parameter estimation techniques* (2nd ed.). New York, NY: Marcel Dekker.

Banilower, E. R., Boyd, S. E., Pasley, J. D., & Weiss, I. R. (2006). *Lessons from a decade of mathematics and science reform: A capstone report for the local systemic change through teacher enhancement initiative.* Horizon Research, Inc. http://www.pdmathsci.net/reports/capstone.pdf

Banilower, E. R., Heck, D. J., & Weiss, I. R. (2007). Can professional development make the vision of the *Standards* a reality? The impact of the national science foundation's local systemic change through teacher enhancement initiative. *Journal of Research in Science Teaching, 44*(3), 375–395. https://doi.org/10.1002/tea.20145

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences, 13*(1), 1–14.

Barab, S., Gresalfi, M., & Ingram-Goble, A. (2010). Transformational play: Using games to position person, content and context. *Educational Researcher, 39*(7), 525–536.

Barab, S., Scott, B., Siyahhan, S., Goldstone, R., Ingram-Goble, A., Zuiker, S. J., & Warren, S. (2009). Transformational play as a curricular scaffold: Using videogames to support science education. *Journal of Science Education and Technology, 18*(4), 305.

Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *ETR&D, 53*(1), 86–107.

Black, L., Williams, J., Hernandez-Martinez, P., Davis, P., Pampaka, M., & Wake, G. (2010). Developing a 'leading identity': The relationship between student's mathematical identities and their career and higher education aspirations. *Educational Studies in Mathematics, 73*(1), 55–72.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*, 77–84.

Bradshaw, L., Izsák, A., Templin, J., & Jacobson, E. (2014). Diagnosing teachers' understandings of rational numbers: Building a multidimensional test with the diagnostic classification framework. *Educational Measurement: Issues and Practice, 33*(1), 2–14.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*(2), 141–178.

Brushlinskii, A. V. (2004). The activity approach and psychology. *Journal of Russian and East European Psychology, 42*(2), 69–81.

Burke, J. P., Orrill, C. H., & Shaffer, D. W. (2012). Epistemic network analysis for exploring connectedness in teacher knowledge. In L. R. Van Zoest, J.-J. Lo, & J. L. Kratky (Eds.), *Proceedings of the Thirty-Fourth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education: Navigating Transitions Along Continuums* (pp. 559). Western Michigan University.

Clements, D. H., Sarama, J., Spitler, M. E., Lange, A. A., & Wolfe, C. B. (2011). Mathematics learned by young children in an intervention based on learning trajectories: A large-sclae cluster randomized trial. *Journal for Research in Mathematics Education, 42*(2), 127–166. https://doi.org/10.5951/jresematheduc.42.2.0127

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Sage Publications Ltd.

Chester, N. C., Ruis, A. R., Collier, W., Swiecki, Z., Arastoopour, G., & Shaffer, D. W. (2015). A novel paradigm for engineering education: Virtual internships with individualized mentoring and assessment of engineering thinking. *Journal of Biomechanical Engineering, 137*(2):024701.

Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education* (pp. 307–333). Erlbaum.

Cobb, P., & Steffe, L. (1983). The constructivist researcher as teacher and model builder. *Journal for Research in Mathematics Education, 14(*2), 83–94. https://doi.org/10.2307/748576

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Cobb, P., Jackson, K., Henrick, E., & Smith, T. M. (2018). *Systems for instructional improvement: Creating coherence from the classroom to the district office*. Harvard Education Press.

Cobb, P., Jackson, K., Smith, T., Sorum, M., & Henrick, E. (2013). Design research with educational systems: Investigating and supporting improvements in the quality of mathematics teaching and learning at scale. In B. J. Fishman, W. R. Penuel, A.R. Allen & B. H. Cheng (Eds.), *Design based implementation research: Theories, methods, and exemplars* (Vol. 112, pp. 320–349). National Society for the Study of Education Yearbook at Teachers College.

Cognition and Technology Group at Vanderbilt (CTVG). (1992). The Jasper Series as an example of anchored instruction: Theory, program description, and assessment data. *Educational Psychologist, 27*(3), 291–315.

Cognition and Technology Group at Vanderbilt (CTVG) (1994). From visual word problems to learning communities: Changing conceptions of cognitive research. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice*. MIT Press/Bradford Books.

Cole, M. (1998). *Cultural psychology: A once and future discipline*. Harvard University Press.

Cole, M., & Engeström, Y. (1993). A cultural-historical approach to distributed cognition. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 1–46). Cambridge University Press.

Cole, M., & Wertsch, J. V. (1996). Beyond the individual-social antinomy in discussions of Piaget and Vygotsky. *Human Development, 39*(5), 250–256.

Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), New directions in educational technology (pp. 15–22). Springer.

Confrey, J., Gianopulos, G., McGowan, W., Shah, M., & Belcher, M. (2017). Scaffolding learner-centered curricular coherence using learning maps and diagnostic assessments designed around mathematics learning trajectories. *ZDM Mathematics Education, 49*, 717–734. https://doi.org/10.1007/s11858-017-0869-1

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th Ed.). Sage.

Davydov, V. V. (1975) The psychological characteristics of the "prenumerical" period of mathematics instruction. In L. P. Steffe (Ed.), *Soviet studies in the psychology of learning and teaching mathematics* (Vol. 7). School Mathematics Study Group.

De Sousa, F., Rasmussen, I., & Pierroux, P. (2018). Zombies and ethical theories: Exploring transformational play as a framework for teaching with videogames. *Learning, Culture and Social Interaction, 19*, 40–50.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Ejersbo, L., Engelhardt, R., Frølunde, L., Hanghøj, T., Magnussen, R., & Misfeldt, M. (2008). Balancing Product Design and Theoretical Insight. In A. Kelly, R. Lesh & J. Baek (Eds.), *The Designer: Journal of the International Society for Design and Development in Education* (pp. 149-163). Lawrence Erlbaum Associates.

Engeström, Y. (1993). Developmental studies of work as a testbench of activity theory: The case of primary care medical practice. In S. Chaiklin & J. Lave (Eds.)*, Understanding practice: Perspectives on activity and context* (pp. 64–103). Cambridge University Press.

Engeström, Y. (1999a). Activity theory and individual and social transformation. *Perspectives on Activity Theory, 19*(38), 19–30.

Engeström, Y. (1999b). Innovative learning in work teams: Analyzing cycles of knowledge creation in practice. *Perspectives on Activity Theory, 377*, 404.

Engeström Y. (2001a) Making expansive decisions: An activity-theoretical study of practitioners building collaborative medical care for children. In C. M. Allwood, & M. Selart (Eds.) *Decision making: Social and creative dimensions*. Springer. https://doi.org/10.1007/978-94-015-9827-9_14

Engeström, Y. (2001b). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work, 14*(1), 133–156.

Engeström, Y. (2004). Learning by expanding: An activity-theoretical approach to developmental research. *Educational Researcher*, 16–25.

Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review, 5*(1), 1–24.

Erlwanger, S. H. (1973). Benny's conception of rules and answers in IPI mathematics. *Journal of Children's Mathematical Behavior, 1*(2), 7–26.

Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S., Doolittle, F., & Warerm E. (2011). *Middle school mathematics professional development impact study: Findings after the second year of implementation*

(NCEE 2011–4024). U.S. Department of Education. https://ies.ed.gov/ncee/pubs/20114024/index.asp

Gee, J. P. (2014). *An introduction to discourse analysis: Theory and method* (4th ed.). Routledge.

Ginsburg, H. P. (1997). In *Entering the child's mind: The clinical interview in psychological research and practice*. Cambridge University Press.

Gresalfi, M., & Barnes, J. (2012). Consequential feedback as a means of supporting student engagement and understanding. In J. van Aalst, K. Thompson, M. J. Jacobson & P. Reimann (Eds.), *The future of learning: Proceedings of the 10th International Conference of the Learning Sciences* (pp. 403–410). International Society of the Learning Sciences.

GSI Teaching & Resource Center (n.d.). Overview of learning theories. Retrieved September 20, 2020, from https://gsi.berkeley.edu/gsi-guide-contents/learning-theory-research/learning-overview/

Hall, R. (2000). Videorecording as theory. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 647–664). Lawrence Erlbaum.

Hegedus, S. J., & Roschelle, J. (2013). *The SimCalc vision and contributions*. Springer.

Hickman, M. (2015). Engaging students with pre-recorded 'live' reflections on problem-solving with *Livescribe* pens. *Research in Mathematics Education, 15*(2), 195–196. https://doi.org/10.1080/14794802.2013.797738

Izsák, A., Orrill, C. H., Cohen, A., & Brown, R. E. (2010). Measuring middle grades teacher's understanding of rational numbers with the mixture Rasch model. *Elementary School Journal, 110*(3), 279–300.

Izsák, A., & Templin, J. (2016). Coordinating conceptualizations of mathematical knowledge with psychometric models: Opportunities and challenges. In A. Izsák, J. T. Remillard, & J. Templin (Eds.), *Psychometric methods in mathematics education: Opportunities, challenges, and interdisciplinary collaborations. Journal for Research in Mathematics Education Monograph, Vol. 15* (pp. 5–30). Reston, VA: National Council of Teachers of Mathematics.

Jaworski, B., & Potari, D. (2009). Bridging the macro-and micro-divide: Using an activity theory model to capture sociocultural complexity in mathematics teaching and its development. *Educational Studies in Mathematics, 72*(2), 219–236.

Johnson, R. B., Onwuegbuzie, A., J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112-133. https://doi.org/10.1177/1558689806298224

Jones, K. (2000). Providing a foundation for deductive reasoning: Student's interpretations when using dynamic geometry software and their evolving mathematical explanations. *Educational Studies in Mathematics, 44*(1), 55–85.

Kim, S., Kwak, M., Cardozo-Gaibisso, L., Buxton, C., & Cohen, A.S. (2017). Statistical and qualitative analyses of students' answers to a constructed response test of science inquiry knowledge. *Journal of Writing Analytics, 1*, 82–102.

Koehler, M. J., & Mishra, P. (2005). What happens when teachers design educational technology? The development of technological pedagogical content knowledge. *Journal of Educational Computing Research, 32*(2), 131–152. https://doi.org/10.2190/0EW7-01WB-BKHL-QDYV

Kosko, K. W., Ferdig, R. E., & Zolfaghari, M. (2020). Preservice teacher's professional noticing when viewing standard and 360 video. *Journal of Teacher Education*, *72*, 84-297. https://doi.org/10.1177/0022487120939544

Kozulin, A. (1996). Vygotsky in context. In A. Kouzlin (Ed.), *Thought and language* (pp. 6–56). The MIT Press.

Kuutti, K. (1996). Activity theory as a potential framework for human-computer interaction research. In B. Nardi (Ed.), *Context and consciousness: Activity theory and human-computer interaction* (pp. 17–44). MIT Press.

Laborde, C. (2005). Robust and soft constructions: Two sides of the use of dynamic geometry environments. In S.-C. Chu, H.-C. Lew, & W.-C. Yang (Eds.), *Proceedings of the 10th Asian Technology Conference in Mathematics* (pp. 22–35). Retrieved from http://epatcm.any2any.us/EP/EP2005/2005P279/fullpaper.pdf

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

Lee, V. R., Drake, J., & Williamson, K. (2015). Let's get physical: K-12 students using wearable devices to obtain and learn about data from physical activities. *Tech Trends, 59*(4), 46–53.

Leont'ev, A. N. (1974). The problem of activity in psychology. *Soviet Psychology, 13*(2), 4–33.

Leont'ev, A. N. (1981). *Problems of the development of the mind*. Progress.

Manizade, A. G., Buchholtz, N., & Beswick, K. (2023). The evolution of research on teaching mathematics: International perspectives in the digital era: Introduction. In A. G. Manizade, N. Buchholtz, & K. Beswick (Eds.,), *The evolution of research on teaching mathematics: International perspectives in the digital era.* Springer.

Martinovic, D., & Manizade, A. G. (2020). Teachers using GeoGebra to visualize and verify conjectures about trapezoids. *Canadian Journal of Science, Mathematics and Technology Education, 20*(3), 485–503.

Maxwell, J. A. (2013). In *Qualitative research design: An interactive approach* (3rd Ed.). Sage.

Mayer, R. E. (2005). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 1–18). Cambridge University Press.

McClain, K. (2002). A methodology of classroom teaching experiments. In S. Goodchild & L. English (Eds.), *Researching mathematics classrooms: A critical examination of methodology* (pp. 91–118). Library of Congress.

McKenney, S., & Reeves, T. (2012). *Conducting educational design research: What it is, how we do it, and why*. Routledge.

McKenney, S., & Reeves, T. C. (2013). Educational design research. In *Handbook of research on educational communications and technology* (pp. 131–140). Springer.

McKenney, S., & Reeves, T. C. (2019). *Conducting educational design research.* Routledge.

Medley, D. M. (1987). Evolution of research on teaching. In M. M. Dunkin (Ed.), *The international encyclopedia of teaching and teacher education* (pp. 105–113). Pergamon.

Nagar, G. G. (2019). *Conceptualization of and reasoning about mathematical structures: Variance and invariance in a dynamic geometry environment.* [Unpublished doctoral dissertation]. University of Massachusetts Dartmouth.

Nagar, G. G., Hegedus, S. J., & Orrill, C. H. (2022). High school teacher's discernment of invariant properties in a dynamic geometry environment. *Educational Studies in Mathematics, 111*(1), 127–145. https://doi.org/10.1007/s10649-022-10144-6

Niess, M. L. (2005). Preparing teachers to teach science and mathematics with technology: Developing a technology pedagogical content knowledge. *Teaching and Teacher Education*, *21*, 509–523. https://doi.org/10.1016/j.tate.2005.03.006

Orrill, C. H., & Cohen, A. (2016). Purpose and conceptualization: Examining assessment development questions through analysis of measures of teacher knowledge. In A. Izsák, J. T. Remillard, & J. Templin (Eds.), *Psychometric methods in mathematics education: Opportunities, challenges, and interdisciplinary collaborations* (pp. 139–153). *Journal for Research in Mathematics Education* Monograph Series No. 15. National Council of Teachers of Mathematics.

Orrill, C. H., & Shaffer, D. W. (2012). Exploring connectedness: Applying ENA to teacher knowledge. In J. van Aalst, K. Thompson, M. J. Jacobson, & P. Reimann (Eds.), *The Future of Learning: Proceedings of the 10th International Conference of the Learning Science (ICLS 2012)—Volume 1, Full papers.* (Vol. 1, pp. 175–179). International Society of the Learning Sciences.

Penuel, W. R., Fishman, B. J., Haugan Cheng, B., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher, 40*(7), 331–337.

Piaget, J., & Cook, M. (1952). *The origins of intelligence in children* (Vol. 8, No. 5, pp. 18). International Universities Press.

Reeves, T. C. (2011). Can educational research be both rigorous and relevant? In *Educational handbook of design research methods in education* (pp. 149–163).

Robins, G. L., (2015). *Doing social network research: Network-based research design for social scientists.* Sage.

Rogoff, B. (2008). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. In K. Hall, P. Murphy & J. Soler (Eds.), *Pedagogy and practice: Culture and identities* (pp. 58- 74). Sage.

Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. *Applied Psychological Measurement, 14*(3), 271–282. https://doi.org/10.1177/014662169001400305

Roth, W. M. (2012). Societal mediation of mathematical cognition and learning. *Orbis Scholae, 6*(2), 7–22.

Rozin, V. M. (2004). Value foundations of conceptions of activity in psychology and contemporary methodology. *Journal of Russian and East European Psychology, 42*(3), 72–89.

Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications.* New York, NY: Guilford.

Ruthven, K. (2018). Constructing dynamic geometry: Insights from a study of teaching practices in English schools. In G. Kaiser, H. Forgasz, M. Graven, A. Kuzniak, E. Simmt, & B. Xu (Eds.), *Invited Lectures from the 13th International Congress on Mathematical Education* (pp. 521–540). Springer.

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences, 231*, 18–36. https://doi.org/10.1080/10508406.2013.778204

Schmittau, J. (2004). Vygotskian theory and mathematics education: Resolving the conceptual-procedual dichotomy. *European Journal of Psychology in Education, 19*, 19–43.

Schmittau, J. (2005). The development of algebraic thinking. *Zentralblatt für Didaktik der Mathematik, 37*, 16–22. https://doi.org/10.1007/BF02655893

Schmittau, J. (2011). The role of theoretical analysis in developing algebraic thinking: A Vygotskian perspective. In J. Cai & E. Knuth (Eds.), *Early algebraization: A global dialogue from multiple perspectives* (pp. 71–85). Springer. https://doi.org/10.1007/BF02655893

Shaffer, D. W. (2004). Epistemic frames and islands of expertise: Learning from infusion experiences. In *Proceedings from the 6th International Conference of the Learning Sciences* (ICLS2004): *Embracing diversity in the learning sciences* (pp. 473–480). Lawrence Erlbaum.

Shaffer, D. W. (2006). Epistemic frames for epistemic games. *Computers and Education, 46*(3), 223–234.

Shaffer, D. W. (2009). Wage the kennel: Games, frames, and the problem of assessment. In R. Ferdig (Ed.), *Handbook of research on effective electronic gaming in education* (Vol. 2, pp. 577–592). IGI Global.

Shaffer, D. W. (2017). *Quantitative ethnography.* Cathcart Press.

Shaffer, D. W. (2018a). Big data for thick description of deep learning. In K. Millis, D. Long, J. Magliano, & K. Weimer (Eds.), *Deep comprehension: Multi-disciplinary approaches to understanding, enhancing, and measuring comprehension* (pp. 262–275). Routledge/Taylor Francis.

Shaffer, D. W. (2018b). Epistemic network analysis: Understanding learning by using big data for thick description. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International handbook of the learning sciences* (pp. 520–531). Routledge.

Shaffer, D. W., & Ruis, A. R. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Grasevic (Eds.), *Handbook of learning analytics* (pp. 175–187). Society for Learning Analytics Research.

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics, 3*(3), 9–45. https://doi.org/10.18608/jla.2016.33.3

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A. A., & Mislevy, R. (2009). Epistemic network analysis: A prototype for 21st century assessment of learning. *International Journal of Learning and Media, 1*(2), 33–53.

Sherin, M. G., Russ, R. S., Sherin, B. L., & Colestack, A. (2008). Professional vision in action: An exploratory study. *Issues in Teacher Education, 17*(2), 27–46.

Simon, M. A., Kara, M., Norton, A., & Placa, N. (2018). Fostering construction of a meaning for multiplication that subsumes whole-number and fraction multiplication: A study of the learning through activity research program. *The Journal of Mathematical Behavior, 52*, 151–173.

Starrett, A., Yow, J., Lotter, C., Irvin, M. J., & Adams, P. (2021). Teachers connecting with rural students and places: A mixed methods analysis. *Teaching and Teacher Education, 97*:103231. https://doi.org/10.1016/j.tate.2020.103231

Steffe, L. P. (1983). The teaching experiment in a constructivist research program. In M. Zweng, T. Green, J. Kilpatrick, H. Pollack, & M. Suydam (Eds.), *Proceedings of the fourth International Congress on Mathematical Education* (pp. 469–471). Birkhauser.

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education* (pp. 267–307). Erlbaum.

Strohmaier, A., MacKay, K., Obersteiner, A., & Reiss, K. (2020). Eye-tracking methodology in mathematics education research: A systematic literature review. *Educational Studies in Mathematics, 104*, 147–200.

Tallman, M. A., & Weber, E. (2015). Toward a framework for attending to reflexivity in the context of conducting teaching experiments. In T. G. Bartell, K. N. Bieda, R. T. Putnam, K. Bradfield, & H. Dominguez (Eds.), Proceedings of the 37thannual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (pp. 1298–1305). East Lansing, MI: Michigan State University.

Tatsuoka, C., Clements, D. H., Sarama, J., Izsák, A., Orrill, C. H., de la Torre, J., … Tatsuoka, K. K. (2016). Developing workable attributes for psychometric models based on the Q-matrix. In A. Izsák, J. T. Remillard, & J. Templin (Eds.), *Psychometric methods in mathematics education: Opportunities, challenges, and interdisciplinary collaborations. Journal for Research in Mathematics Education Monograph, Vol. 15* (pp. 73–96). Reston, VA: National Council of Teachers of Mathematics.

van der Kleij, F., Adie, L., & Cumming, J. (2019). Feasibility and value of using a GoPro camera and iPad to study teacher-student assessment feedback interactions. In B. P. Veldkemp & C. Sluijter (Eds.), *Theoretical and practical advances in computer-based educational measurement* (pp. 339–359). SpringerOpen.

van Gog, T. & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*, 95–99.

Wagner, E., Orme, S., Turner, H., & Yopp, D. (2017). Encouraging example generation: A teaching experiment in first-semester calculus. *PRIMUS, 27*(2), 212–234. https://doi.org/10.1080/10511970.2016.1194340

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher, 37*(8), 469–479.

Weiland, T., Orrill, C. H., Nagar, G. G., Brown, R. E., & Burke, J. (2020). Framing a robust understanding of proportional reasoning for teachers. *Journal of Mathematics Teacher Education, 24*(5), 179–202.https://doi.org/10.1007/s10857-019-09453-0

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.

Wooldridge, A. R., Carayon, P., Shaffer, D. W., & Eagan, B. (2018). Quantifying the qualitative with epistemic network analysis: A human factors case study of task-allocation communication in a primary care team. *IISE Transactions on Healthcare Systems Engineering, 8*(1), 72–82. https://doi.org/10.1080/24725579.2017.1418769

Yamagata-Lynch, L. C. (2003). Using activity theory as an analytic lens for examining technology professional development in schools. *Mind, Culture, and Activity, 10*(2), 100–119.

Yamagata-Lynch, L. C. (2007). Confronting analytical dilemmas for understanding complex human interactions in design-based research from a cultural—Historical activity theory (CHAT) framework. *The Journal of the Learning Sciences, 16*(4), 451–484.

Yamagata-Lynch, L. C. (2010). *Activity systems analysis methods: Understanding complex learning environments*. Springer Science & Business Media.

Yoon, K. S., Duncan, T., Lee, S. W-Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (REL 2007-No. 33). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from http://ies.ed.gov/ncee/edlabs

Zolfaghari, M., Austin, C. K., Kosko, K. W., & Ferdig, R. (2020). Creating asynchronous virtual field experiences with 360 video. *Journal of Technology and Teacher Education, 28*(2), 315–320.