

INTRODUCTORY STATISTICS: PREPARING IN-SERVICE MIDDLE-LEVEL MATHEMATICS TEACHERS FOR CLASSROOM RESEARCH

JENNIFER L. GREEN

Montana State University

jgreen@montana.edu

WENDY M. SMITH

University of Nebraska - Lincoln

wsmith5@unl.edu

APRIL T. KERBY

Winona State University

akerby@winona.edu

ERIN E. BLANKENSHIP

University of Nebraska - Lincoln

erin.blankenship@unl.edu

KENDRA K. SCHMID

University of Nebraska Medical Center

kkschmid@unmc.edu

MARY ALICE CARLSON

Montana State University

mary.carlson5@montana.edu

ABSTRACT

In this study, we examined how in-service middle-level mathematics teachers used statistics in their own classroom research. Using an embedded single-case design, we analyzed a purposefully selected sample of nine teachers' classroom research papers, identifying several themes within each phase of the statistical problem solving process to summarize how teachers 1) planned studies and collected data, 2) analyzed data, and 3) interpreted results. The results illustrate the varying ways in which teachers used statistics to make data-based decisions about their classrooms, revealing teachers' early development in their statistical thinking and suggesting that teachers' required knowledge of statistics is multi-faceted, requiring both a pedagogical component and statistical knowledge for the teaching profession. Such findings have important implications for how we, as teacher educators, can best meet teachers' professional needs.

Keywords: *Statistics education research; Data-driven decision making; Teacher development; Statistical thinking*

1. INTRODUCTION

There is a growing emphasis on using data to make informed educational decisions (Ikemoto & Marsh, 2007; Marsh, Payne, & Hamilton, 2006). This, coupled with sweeping changes in statistics curricula across several countries (Batanero, Burrill, & Reading, 2011) place additional instructional and professional demands on teachers. Teachers must understand the statistical concepts they teach (Common Core State Standards Initiative (CCSSI), 2010; Conference Board of the Mathematical

Sciences (CBMS), 2001, 2012; Franklin et al., 2015), as well as the statistics they encounter and are expected to use as professionals (Bargagliotti, 2014; Mandinach & Gummer, 2012; Scheaffer & Jacobbe, 2014). In addition to teaching students how to think and reason with data, teachers must collect, summarize and interpret data in meaningful ways that allow them to make decisions about curriculum, instruction, and teaching practices (Hamilton et al., 2009; Mandinach, 2012).

Teachers have access to many potentially valuable sources of information about their students, such as daily interactions with students, homework assignments, unit exams, and standardized tests. However, many teachers have limited statistical preparation for how to summarize and interpret that information to make informed decisions about how to improve student learning (Mandinach, 2012; Means, Chen, DeBarger, & Padilla, 2011). Additionally, it can be challenging for teachers to recognize when it is appropriate to use quantitative or qualitative methods, and to know how best to do so. These challenges, coupled with the task of understanding and teaching various statistical concepts to students, further expose the need for us, as teacher educators, to understand how teachers use statistical concepts to assess students and their own practice. This broader understanding of teachers' uses of statistics can help to enrich the ways in which we support teachers' statistical development and growth. With statistics playing an increasing role in mathematics teachers' instructional (CBMS, 2012; CCSSI, 2010) and professional (Mandinach, 2012) responsibilities, developing valuable learning opportunities is critical for effectively preparing teachers and supporting their continued growth and development in statistics.

Existing studies primarily focus on understanding how teachers develop as instructors of statistics, emphasizing teacher preparation for how to *teach* statistics (Batanero et al., 2011; Franklin, 2013). However, it is equally important to understand how teachers use statistics to understand and make data-based decisions about their practice. Thus, more research into preparation for *using* statistics as a teaching professional is needed. Action research, a process in which “teachers identify a problem within their own practice, devise a plan to study it, systematically collect and analyze data, and use the findings to inform practice” (Smith & Heaton, 2013, p. 148), provides a context for educational researchers and teacher educators to understand how teachers gather, interpret, and use data in their own work. Although traditionally explored qualitatively, action research projects can also be informed by quantitative approaches (Chen, Huang, & Zeng, 2017).

The goal of this research was to investigate how in-service middle-level mathematics teachers used statistics to study the impact of changes in their own classroom practice on their teaching and their students. Our research was guided by the following research questions:

- 1) After having taken a course in introductory statistics, how and to what extent do teachers use statistics in an action research project?
- 2) What do teachers' uses of quantitative data in action research projects reveal about their development in statistical thinking?

2. THEORETICAL PERSPECTIVE

Our study draws on and makes use of three investigative processes that support *practitioner inquiry*: action research, data-driven decision making (DDDM), and statistical problem solving. We view the three processes as nested, with the outer layers creating the context and need for those contained within.

In practitioner inquiry, the teacher is not just an instructor teaching the class, but “simultaneously a researcher who is continuously engaged in inquiry with the ultimate purpose of enriching students' learning and life chances” (Cochran-Smith & Lytle, 2009, p. vii). Action research, the first investigative process, is a form of practitioner inquiry common in teacher education and professional development settings (e.g., Zeichner, 2001). It is a systematic problem-solving process that teachers use to understand and improve their own practice. When engaged in action research, teachers identify problems of practice they want to understand, study changes in their own classroom practices, and examine how those changes impact their teaching and their students. During action research, teachers' problems of practice are traditionally explored qualitatively, but the problems can also be informed by quantitative approaches (Chen et al., 2017). For example, teachers could explore students' attitudes towards mathematics both qualitatively and quantitatively through individual student interviews and surveys, giving a more holistic picture of outcomes stemming from changes in teaching practice. The purpose



Figure 1. Theoretical perspectives informing our study

of action research is not to generalize conclusions to a broader population, but to integrate theory and practice (Vaughan & Burnaford, 2016), to inform instructional decisions, and to improve student outcomes (Cochran-Smith & Lytle, 1993; Cochran-Smith & Lytle, 2009; Mills, 2010).

DDDM, the second investigative process, is a form of action research: it places emphasis on transforming data into “actionable knowledge” (Marsh, 2012, p. 3) to make and evaluate changes in district, school, or teacher practice. Although both qualitative and quantitative data can inform instructional decisions, DDDM tends to emphasize the use of quantitative sources, especially as schools respond to accountability policies (Marsh et al., 2006). We situate DDDM within action research with an important caveat. As Cochran-Smith and Lytle (2009) have pointed out, the current accountability climate, to which DDDM at least in part responds, constrains practitioner inquiry when it focuses on the means rather than the ends of education, emphasizes transmission models of teaching and learning, and makes teachers the recipients of other peoples’ knowledge. When DDDM supports teachers in seeking deeper understanding of their own teaching and of students, gives teachers tools to push back against overly simplified views of “what works” in education, and reveals rather than hides the nuances of instructional decisions, it is rightly conceived as action research.

Although the mantra “data-driven decision making” has been ringing out in schools for many years, the presence of data does not guarantee its effective use (Marsh et al., 2006). To use data effectively, teachers need to become adept statistical problem solvers. We adopt the four-phase statistical problem solving process outlined in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report* (Franklin et al., 2007) as a third investigative process that informs our study. When solving statistical problems, teachers formulate questions, collect data, analyze data, and interpret results, paying close attention to the role of variability throughout the entire process. Understanding and effectively carrying out the statistical problem solving process in a cyclic manner supports the effective use of data as a part of practitioner inquiry.

3. LITERATURE REVIEW

In order to respond to the call to be data-driven decision makers, teachers must think and reason statistically when making decisions that impact the educational opportunities given to their students. The role of statistics in the K–12 curriculum is to create statistically competent citizens (Franklin et al., 2007), but educators must not only support their students in becoming statistically literate; they must also be statistically literate themselves. This can be accomplished by nurturing statistical thinking and reasoning, and teaching through the use of the statistical problem solving process.

3.1. STATISTICAL LITERACY, REASONING, AND THINKING

Franklin et al. (2007) refer to statistical literacy as “the ultimate goal” (p. 1) of K–12 statistics education. In this sense, statistical literacy provides an ends-in-view perspective on statistics education and shines a spotlight on “critical statistical survival skills of both school students and adults” (Shaughnessey, 2007, p. 961). To be statistically literate, individuals must understand the language and tools of the discipline, know what statistical terms mean, understand the use of statistical symbols, and understand and interpret representations of data (Ben-Zvi & Garfield, 2004; Rumsey, 2002).

Although much attention has been given to developing statistical literacy of students in K-12 settings, teachers must also develop as statistically literate adults, who can collect, use, and interpret data to positively impact their own professional communities. Gal (2004) identified two interrelated components that comprise statistical literacy in adults: the ability to “*interpret and critically evaluate* statistical information” and the ability to “*discuss or communicate* their reactions to such statistical information” (p. 49). These abilities require statistical knowledge and dispositions that align with K–12 content and standards (CCSSI, 2010). Teachers must understand elementary probability and graphical displays, as well as more sophisticated concepts such as the fundamental necessity of data (Gal, 2004). In addition, they need to willingly question spurious data-based arguments (Gal, 2004). Clearly if teachers can develop statistical literacy, they in turn are better positioned to help students develop it.

Statistical reasoning and statistical thinking are also critical aspects of developing K-12 students’ and teachers’ statistical literacy. These skills help people make sense of statistical information (Garfield, 2002) and understand why and how statistical investigations are conducted (Chance, 2002; Garfield & delMas, 2010). Together, the knowledge and dispositions required to statistically reason and think with data are critical components of statistical literacy and, ultimately, statistical proficiency in which one is able to successfully connect and use the tools and skills of the discipline.

Several authors have laid out models of statistical proficiency. Garfield’s (2002) Model of Statistical Reasoning identifies five levels of proficiency, providing a trajectory for statistical thinking, but it does not provide directions for how to advance knowledge and skills to achieve higher levels of proficiency. In contrast, Wild and Pfannkuch’s (1999) framework for statistical thinking is comprised of four dimensions that break down the skills needed to develop high levels of proficiency with statistical literacy. These dimensions include investigation and problem solving, types of knowledge and thinking fundamental to statistical reasoning, interrogation, and dispositions for inquiry.

Studies of teachers’ uses of data to inform their own instruction suggest that teachers may face significant gaps in both the knowledge and dispositional components of statistical literacy, reasoning and thinking. For example, a study of teachers’ thinking about data conducted by the U.S. Department of Education (Means et al., 2011) found that teachers struggled to make valid inferences, especially when making sense of differences or trends and “appeared to lose track of what they were trying to figure out” (p. 61) when analyzing scenarios involving grade- or school-level data. Although collaborative teams and strong leadership structures support better use of data in schools (e.g., Lachat & Smith, 2005), more research is needed on defining and developing statistical literacy, thinking and reasoning for teachers. The current study takes up this gap in the literature, studying how middle-level mathematics teachers who completed a graduate course in statistics applied statistical thinking and reasoning skills to local data.

3.2. STATISTICAL PROBLEM SOLVING

Statistical problem solving manifests as the application of statistical thinking and reasoning to particular problems. As defined in the literature, the statistical problem solving process consists of between four and six stages, including specifying the problem, planning, collecting data, analyzing the data, and making conclusions from the data (Brand-Gruwel, Wopereis, & Vermetten, 2005; Mariott, Davies, & Gibson, 2009; Stuart, 1995; Wild & Pfannkuch, 1999). However, students following the same process for the same problem does not guarantee the same results. Garfield (1995) noted that, “Students appear to understand and reconstruct a problem in different ways, leading them to apply different strategies to solve them” (p. 28) and suggested general principles of learning statistics: students learn by constructing knowledge, actively engaging in learning activities, and practicing what we want

them to do well. These principles point to the fact that sound statistical thinking cannot be separated from statistical problem solving (Wild & Pfannkuch, 1999); it is foundational and should be present throughout the entire problem-solving process.

A problem-solving process can be a benefit to both teachers and students (Rossman, Medina, & Chance, 2006) because problem solving gives context to statistical issues and sets up questions whose answers require using data and statistical analyses (Stuart, 1995). However, most elementary statistics courses focus mainly on the analysis phase of the process and neglect the process as a whole (Stuart, 1995). This curricular omission is compounded when teachers are not confident about teaching statistical problem solving, have a negative opinion regarding the course work and are pressed for time [so may leave out the topic or gloss over it], and/or do not fully understand the importance of the problem-solving process. Problem specification, particularly using actual data and situations, is too often under-valued and at best implicitly taught; yet, this crucial skill in statistical literacy is key to all statistical problem solving. Additionally, appropriate assessments for determining whether a student has mastered the problem-solving process need to focus more on the thinking process, the links between the stages, and the fact there may be more than one possible [correct] solution rather than merely focusing on the mechanics of the analysis (Marriott et al., 2009).

The importance of the statistical problem solving process has been reiterated to teachers in documents such as the National Council of Teachers of Mathematics (2000) *Principles and Standards for School Mathematics* which describes key student learning outcomes and the *GAISE Report* (Franklin et al., 2007) which provides a framework for using the statistical problem solving process and the levels of understanding through which students should progress. Thus, effective professional development for teachers needs to include explicit attention to the full cycle of statistical problem solving, from problem specification to the interpretation of findings.

3.3. TEACHER PREPARATION IN STATISTICS

Subject-specific knowledge of content and pedagogy are needed if teachers are to prepare their students to use statistics in the workplace, in their personal lives, and as citizens (Batanero et al., 2011). However, most teachers are not adequately trained to prepare students to become statistically literate citizens (Batanero et al., 2011; Froelich, Kliemann, & Thompson, 2008). Many states do not offer certification in statistics as a subject (Franklin et al., 2007), and most teacher preparation programs include at most one general statistics course. Yet, teachers need to be trained in statistics-specific topics such as designing studies, analyzing data, and using appropriate statistical software, in addition to the pedagogy related to teaching these topics.

Professional development programs that aid teachers in obtaining the knowledge needed to teach statistics also promote teachers' knowledge in using data and statistics themselves (Kazak & Confrey, 2004). For example, the professional development sequence by Makar and Confrey (2004) was created under the assumption that if mathematics teachers are immersed in content beyond the level of what they teach, and developed through their own investigations as statisticians with a context that they find compelling and useful, then they will teach statistics more authentically and their increased content knowledge will translate into improved practice (p. 357).

Makar and Confrey (2005) went one step further in noting, "that teachers themselves need to learn statistical concepts in an environment much like the one recommended for students" (p. 30). Batanero et al. (2011) suggested that professional development for teachers should promote their statistical thinking, work with real data through projects and statistical investigations, allow opportunities to work with technology, and support teachers in making connections to their own practice.

Not only do mathematics teachers need a solid background in statistics in order to teach the content, but they also need to be able to use statistics to interpret student test scores and assess whether pedagogical changes were successful. It is becoming increasingly important for teachers to make instructional decisions based upon their students' performance (Kazak & Confrey, 2004; Makar & Confrey, 2004). Ikemoto and Marsh (2007) encouraged the use of an inquiry-focused data-driven decision making process to make continuous improvements over time. Teachers need to formulate questions regarding curriculum or student improvement they want to answer and a solid understanding of the data at their disposal. After data have been collected and analyzed, teachers should take appropriate actions that are based on the results. Thus, professional development programs are needed

that allow teachers to develop data analysis skills and work with relevant data to support their understanding and motivation to learn statistics, which in turn may help to further their understanding about testing (Ikemoto & Marsh, 2007; Kazak & Confrey, 2004).

Statistics can provide useful tools for teachers conducting their own classroom research, creating opportunities to better understand and support teachers' development as users of statistics. Action research improves outcomes for students, develops context-specific solutions to problems, provides effective professional development, and helps to sustain improvements in teaching and learning (Lai & Robinson, 2006). It is "a way to empower...teachers to be more reflective about their own teaching and engage in a cycle of continuous inquiry, self-evaluation, and improvement" (Miller, 2017, p. 34).

Overall, statistical literacy is a crucial skill, particularly for citizens in a democracy. Most statistical content is taught by mathematics teachers at the elementary and secondary levels, but such teachers often have little to no preparation in teaching statistical literacy. In addition to understanding the statistical problem solving process—from problem specification to modeling to interpreting results—teachers in today's accountability climate are also in the position to be consumers of data via statistical models. Teacher preparation programs do not typically include statistics courses for teachers, so professional development for in-service teachers seems to be the most viable route for improving teachers' statistical literacy. Research is needed to understand the extent to which professional development can support teachers in becoming first statistically literate (as learners) and then proficient in developing statistically literate students.

4. CONTEXT

The Mathematical Education of Teachers 2 (MET2) report (CBMS, 2012) makes the point that teacher preparation programs at best are producing "well-started beginning teachers" (p. 18). Teachers need ongoing, high-quality longitudinal professional development to continue to deepen their knowledge of mathematics for teaching, their knowledge of pedagogy, and their knowledge of student cognitive development. Various professional development opportunities offer teachers such additional instruction and learning experiences in mathematics and statistics (e.g., Advisory Committee on Mathematics Education, 2002; National Centre for Excellence in the Teaching of Mathematics, 2011; University of Wisconsin-Madison, 2017). At the University of Nebraska-Lincoln, Nebraska middle-level teachers have been offered such opportunities to further their education in mathematics and statistics through the Math in the Middle Institute Partnership. In this section, we briefly describe this master's degree program, its statistics course, and its action research projects.

4.1. MATH IN THE MIDDLE INSTITUTE PARTNERSHIP

The Math in the Middle (M^2) Institute Partnership is a master's degree program for middle-level mathematics teachers funded by a grant from the National Science Foundation. The program was designed to span three summers and the two intervening academic years. Teachers completing the 36-credit hour program earn a Master of Arts for Teachers (MAT) degree from the Department of Mathematics or a Master of Arts (MA) degree from the Department of Teaching, Learning and Teacher Education through a combination of on-campus summer courses and academic year hybrid courses. Six cohorts of teachers began the M^2 Institute from 2004 to 2009, and local funding sources have sponsored three additional cohorts that began in 2013 (one cohort) and 2014 (two cohorts).

The program includes seven mathematics courses, one statistics course, three education/pedagogy courses, and one capstone course (see Table 1). Summer courses either run for one week (8am-5pm with daily homework and a substantial end-of-course problem set) or two courses are paired and run for two weeks, Monday through Friday, with one course meeting in the morning and the other in the afternoon. Academic year courses meet for 2–3 days across the semester, with most assignments completed online in between face-to-face meetings.

Most cohorts of teachers have 30–35 participants. Although the focus of the M^2 Institute is on grades 5–8, many of these teachers teach mathematics to multiple grades of students (e.g., K–6, 3–6, 7–8, 7–12) due to the preponderance of rural school districts in Nebraska. A total of 158 teachers earned master's degrees through the M^2 Institute from 2005 to 2011, representing a 98% retention rate to graduation for the six NSF-funded cohorts.

Table 1. Math in the Middle Institute courses

Subject	Course Title
Mathematics	Mathematics as a Second Language
	Functions, Algebra and Geometry for Middle-Level Teachers
	Experimentation, Conjecture and Reasoning
	Discrete Mathematics for Middle-Level Teachers
	Number Theory and Cryptology for Middle-Level Teachers
	Using Mathematics to Understand Our World
Education	Concepts of Calculus for Middle-Level Teachers
	Inquiry into Teaching and Learning
	Curriculum Inquiry
Statistics	Teacher as Scholarly Practitioner
	Statistics for Middle-Level Teachers
Mathematics education	Integrating the Teaching and Learning of Mathematics (Capstone)

More information about the M² Institute courses can be found on the M² Institute website (<http://scimath.unl.edu/MIM/>) and in previous publications about the M² Institute (Heaton, Lewis, Homp, Dunbar, & Smith, 2013; Heaton, Lewis, & Smith, 2009; Heaton, Lewis, & Smith, 2013).

4.2. STATISTICS FOR MIDDLE-LEVEL TEACHERS

The M² Institute statistics course, *Statistics for Middle-Level Teachers* (STAT-MLT), is designed to help teachers improve their statistical reasoning and thinking skills so they can help their students do the same. Teachers typically take the course during their second summer in the program, and the course is most often taught during a one-week timeframe with a follow-up end-of-course project. The topics covered in the course are very similar to those found in a traditional, algebra-based introductory statistics course: data collection and study design, descriptive statistics and graphical representations, one- and two-sample inference procedures (hypothesis tests and confidence intervals), and simple linear regression. However, unlike the traditional algebra-based introductory course which primarily has the goal of creating quantitatively literate citizens and consumers, the STAT-MLT course has two overarching goals: prepare teachers for the statistical content they will teach to their own students and provide teachers with statistical tools and skills they need to understand types of research they might encounter or even conduct as teaching professionals. The topics covered in the course and their presentation are all framed in the context of one of these two goals.

To support the class participants as teachers of statistics, course topics that are represented in the Common Core State Standards for Mathematics (CCSSI, 2010) and /or the Nebraska state standards are covered in more detail than they would be in a traditional introductory course, and emphasis is placed on misconceptions they might encounter from their students. For example, considerable time is devoted to creating and interpreting graphical summaries. Although this is a common topic in the traditional class, as graphing is a topic in numerous grade levels, more time is spent reviewing both good and bad graphs for teachers to gain a better understanding of how graphs can be misleading. In addition, most of the activities used in this class to introduce or reinforce concepts are activities that could be easily adapted to many grade levels and statistical content areas. Also, as part of the end-of-course project, participants are asked to create a lesson or unit plan that reinforces a concept covered in the course. This not only provides the teachers with a lesson they can take straight into their own classroom, but also illustrates to them how statistics concepts can be integrated into their curriculum.

To support the class participants as users of statistics, course topics such as one- and two-sample inference are covered in a way that illustrates how middle-level teachers can use them as practicing statisticians to make informed decisions about their teaching practice and analyze data resulting from their action research projects (see Section 4.3). For example, after we have covered two-sample hypothesis tests and confidence intervals for means, the participants are asked to carry out a data analysis project in which they are provided a set of (fictional) student achievement data linked to teachers who either have or have not participated in a professional development program. We ask the participants to see what they can learn about whether or not the professional development program was effective, but with very little other direction. In groups, the participants then use whatever procedures

they wish to explore the data set and make conclusions. Teachers are encouraged to use Excel or other technology tools to which they have access. They then present their findings to the rest of the class. Our goal with this activity is not only to allow the participants to practice inference procedures, but also to help them realize that there are multiple, valid ways to approach a data analysis. Other examples are concepts presented to the participants where concerted efforts are made to tie course topics to their reference point as teachers. For example, when covering z-scores, we connect these ideas to standardized test scores they may receive for their students, and when discussing the concepts of lurking variables, we present excerpts from the book *Freakonomics* (Levitt & Dubner, 2009) that discuss student achievement and underlying drivers of achievement.

Cognizant of the fact that the participating teachers are engaged in educational research projects (both through conducting their own research and through reading scholarly research articles), we intentionally use teachers' course evaluations and our own ongoing course self-assessments to reevaluate what teachers really need to understand about statistics. Based on our analyses and reflections, we redesign each new course offering accordingly. Since its first offering in July, 2005, the course has evolved and continues to change (Schmid, Blankenship, Kerby, Green, & Smith, 2014), incorporating new aspects such as the aforementioned data analysis project, so that we can help teachers develop not only as teachers of statistics, but also as teaching professionals and classroom researchers who use statistics to make informed decisions about teaching and students.

4.3. ACTION RESEARCH PROJECTS

Teachers plan and carry out action research projects as a final component of their degree. Throughout the entire process, teachers have faculty advisors who help with the design, implementation, analysis, and final write-up of the action research projects. During the fall *Teacher as Scholarly Practitioner* course, there is a team of instructors who helps teachers design the projects. In the spring, each teacher is assigned a faculty advisor to guide the rest of the action research process, including the written reports. Because of the relatively large number of teachers compared to the number of mathematics educators, some teachers are assigned advisors with no background in mathematics education; none are assigned to faculty advisors with strong statistical backgrounds. However, this is due in part to the way the action research projects evolve. With data collection focused on teachers keeping researcher journals, interviewing students, and analyzing student work, they are almost always advised by qualitative researchers.

During this year-long process, teachers begin by identifying a “problem of practice”—something they would like to change about their teaching. Teachers then craft research questions and identify research variables for their problems of practice. Because teachers are studying their own teaching practices in a single, relatively small mathematics class (4–32 students/class) over a relatively short period of time (approximately three months), their inquiries lend themselves to qualitative approaches.

Teachers then develop inquiry plans that include final drafts of their problem statements, purpose statements, research questions, research variables, methods of data collection, and data collection instruments. Teachers are required to collect at least three different types of qualitative data related to each of their two to four research questions. In addition, teachers are required to keep a weekly personal teacher journal during data collection, to collect and analyze student work each week, and to interview students at least once during the projects. Some teachers also choose to survey students or to collect student grades, but most ultimately find that the required forms of data collection provide the best evidence to support their findings. As they carry out their inquiry plans, teachers are asked to periodically analyze their collected data, form assertions, and make adjustments to data collection (such as adding interview questions to explore emerging themes from these intermediate analyses) as needed.

Finally, teachers write reports of their action research projects. They are expected to make at least one assertion per research question, and to support each assertion with at least three pieces of evidence from different sources. Such evidence typically comes from quotations from participant journals, quotations from student interviews, and copies or portions of student work. Teachers are also asked to relate their findings to those from related research literature in their conclusion, and discuss how their action research projects will inform their future teaching practices and how their results will be shared with the broader teaching community.

When teachers collect quantitative data, such as student grades or student survey responses, they are encouraged to report the data using descriptive statistics. For instance, most surveys administered investigate student attitudes using Likert-scale items (strongly disagree to strongly agree). Teachers are encouraged to calculate median and mean responses to each item, along with standard deviations, and also to create frequency graphs for certain survey items to show the changes in the distribution of student responses across the period of data collection. Teachers are also encouraged to find graphical representations for student grade data and other quantitative data.

5. METHODS

We used an embedded, single-case design (Yin, 2009) to investigate how in-service middle-level mathematics teachers use statistics in their classroom research. The M² Institute was our case and teachers' written action research papers resulting from the *Teacher as Scholarly Practitioner* course were our embedded units of analysis. These action research papers provide rich evidence of teachers' abilities to translate what they have learned in the STAT-MLT course to their own research, and help to illustrate how teachers use statistics to make classroom-based decisions.

5.1. SAMPLE

Our data set includes action research papers from the first four cohorts of teachers who completed the M² Institute ($n = 123$). To identify papers for closer analysis, we began by identifying those projects that used statistics in any form ($n = 111$). We then sorted the papers into three groups according to the role statistical analysis played in investigating their overarching problem of practice. If statistical analysis played a central role in answering the teacher's research questions, we placed the project in the "primary" group. Some teachers used statistics alongside qualitative approaches, with quantitative and qualitative analyses making equal contributions to the paper. We labeled this group "complementary." Finally, in some projects, statistics played a small role in an otherwise qualitative investigation. We labeled this group "supporting."

Case study research is useful in illuminating both *typical* and *revelatory* examples of the phenomena under study (Yin, 2009). After sorting papers into three groups, we identified a subset of papers that were typical, in that they exemplified uses of features typical to the majority of papers in that group. We also identified a subset of papers that were revelatory, in that they represented a teacher's use of statistics that, although atypical, was worth investigating in greater depth. This process left us with a purposefully selected subset of action research papers ($n = 5$ typical and 4 revelatory) that represents the variety of statistical analyses completed by teachers in each of the first four cohorts. Figure 2 shows our process for selecting papers for close analysis.

5.2. ANALYSIS

We engaged in a three-part analysis process led by two authors who had collaboratively taught the STAT-MLT course. First, both authors independently read each paper and documented the ways teachers used and reported statistics. Within this process, the authors deliberately looked for how teachers applied ideas and topics from STAT-MLT to action research. Because both authors had collaboratively taught all of the first four iterations of the STAT-MLT course, they were familiar with the course content and were able to recognize when and how teachers used this content for their analyses. In particular, the authors looked for instances when teachers referenced statistical concepts the course had emphasized, such as data collection, descriptive statistics and graphical representations, one- and two-sample inference procedures (hypothesis testing and interval estimation), and regression.

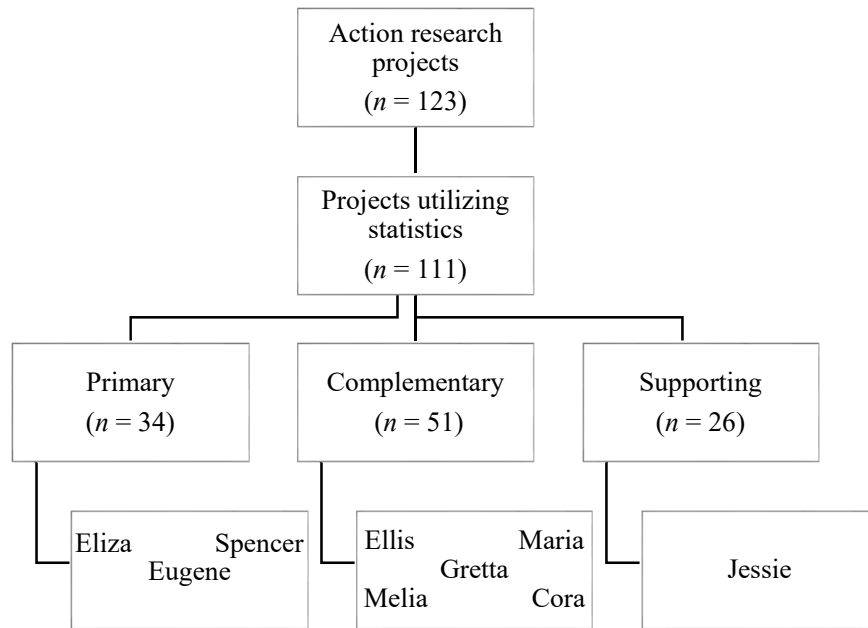


Figure 2. Diagram of sampling process

Next, the authors identified and noted any limitations and/or successes teachers faced when employing such methods. For example, both noted that one teacher wrote about using randomization to interview students in random order, but the teacher did not explain how that order was determined randomly instead of haphazardly. As illustrated in Table 2, notes such as these were organized in a table with rows identifying the paper analyzed, and columns identifying the type of recorded information: 1) “Statistical Methods Used,” 2) “Manner of Use,” and 3) “Limitations and/or Successes Employing the Statistical Methods.” The researchers also recorded the page number(s) on which they made the corresponding observation(s). Because a teacher could use multiple statistical methods in a single action research project, each paper was allowed to have multiple table entries.

Table 2. Subset of table entries one author created for the action research papers analyzed

Paper ID	Statistical methods used	Manner of use	Limitations and/or successes employing the statistical methods
1	Randomization	Used randomization to interview students in random order	No explanation of how order was determined randomly instead of haphazardly
1	Numerical summaries	Calculated and reported mean differences or difference in two means	No reported standard deviation(s) and no description of how differences calculated
2	Randomization	Randomly selected students to interview	Describes how the sample was chosen by drawing six student names from a hat

In phase 3, authors identified when in Franklin et al.’s (2007) statistical problem solving process the teachers used statistics. Categories represented key aspects of the statistical problem solving process and provided a framework for further analyzing teachers’ use of statistics. Using the labels planning studies and collecting data; analyzing data; and interpreting results, the authors categorized their table entries to indicate which stage(s) of the statistical problem solving process each entry represented. They examined entries within each category for commonalities and identified emerging themes that described the teachers’ uses of statistics across the phases of the statistical problem solving process. As they identified themes, the authors looked for instances that supported these themes and for those that did

not, being sure to reference the raw data throughout the validation process. Although the level of statistical sophistication and rigor varied across the action research papers sampled, distinct uses of statistics in each of the three stages of the statistical problem solving process recurred throughout the projects, validating the initial categories and themes identified.

6. RESULTS

Across the action research projects, teachers used statistics for a variety of purposes throughout the different phases of the statistical problem solving process (see Table 3). For example, when planning their studies and collecting data, teachers used their knowledge of statistics to collect longitudinal data, identify which data would help answer their research questions of interest, and incorporate randomization in the data collection process. When analyzing the data, teachers used statistics when determining which levels of analysis to use for their data, and when summarizing their data graphically and numerically. As they interpreted their results, teachers also encountered statistics when determining the scope of inference and stating potential limitations and alternative explanations for their studies. Because of the nature of study design and implementation, all of the action research projects had either explicit or implicit evidence to support each of the themes except one: incorporating randomization. Among the nine papers examined closely, six of them reported using randomization.

We did not identify teachers using statistics or engaging in statistical reasoning to formulate questions (phase 1 in the GAISE framework). Although we think teachers likely did “clarify the problem at hand” and “formulate one (or more) questions that can be answered with data” (Franklin et al., 2007, p. 11), statistics educators were not involved in the action research process. Thus, we only had access to processes that teachers documented in their final papers. In the following sections, we discuss in more depth the primary ways in which teachers used statistics during action research and what these uses reveal about teachers’ development in statistical thinking based on Wild and Pfannkuch’s (1999) four-dimensional model of statistical thinking.

Table 3. When and why teachers use statistics during action research within the statistical problem solving process

Phase	Purpose
Planning studies and collecting data	Collect longitudinal data
	Identify data to answer a research question
	Incorporate randomization
Analyzing data	Determine levels of analysis
	Summarize data graphically and numerically
Interpreting results	Determine scope of inference
	State potential limitations and alternative explanations

6.1. PLANNING STUDIES AND COLLECTING DATA

Teachers’ use of statistics to collect longitudinal data Action research projects often center on teachers making a specific change to their teaching practices and then exploring what happens to targeted student outcomes. In order to explore potential changes in student outcomes, teachers need to have multiple measurements of the targeted outcomes over time. This type of research, at a minimum, requires that outcomes be measured both before and after a change has been made.

In the action research projects we analyzed, there were several instances when teachers designed studies to investigate how their students changed over time (if it all). The teachers primarily used pre- and poststudy data to explore changes in student outcomes, but they implemented their studies in different ways. In some instances, teachers collected baseline data prior to making changes in their practice, and used these data to later compare how students changed over time. One teacher, Jessie, explicitly acknowledged the importance of collecting such baseline data, stating that “baseline data was necessary to collect early in the study before students had had experience with working in homework teams.”

The pre- and poststudy data teachers used were usually, although not always, collected with the same instrument at each measurement occasion. For example, Cora noted, “The second survey included many of the same questions from the presurvey, as well as more detailed ones.” In this instance, the postsurvey included all but one item from the presurvey as well as five additional questions. Cora did not explain why the postsurvey was different from the presurvey.

Teachers also made decisions about how to track student outcomes over time, with some prioritizing students’ anonymity over the ability to match pre-post results. For instance, Cora explained, “Putting a name on the survey [pre and post] was optional to get an honest response.” In these instances, classroom-level changes were reported, whereas in other instances, student-level changes were explored.

Teachers’ use of statistics to identify data to answer a research question As part of the action research project, teachers were required to create two to four research questions of interest, as well as collect at least three different types of data related to each of these questions. Most of the data were qualitative, but some teachers chose to collect quantitative data to help answer some of their research questions. Five of the nine papers closely studied included qualitative and quantitative data collection, with both playing a balanced and complementary role in providing evidence for assertions related to the research questions (see Figure 2).

By incorporating a quantitative component to their research, teachers had to identify not only what data to collect, but also which data would best answer a research question. For example, when exploring the “accuracy of students’ assessments of their own understanding,” Gretta identified “the best data for this research question was the survey.” She “also used the students’ assessments [tests] and homework as evidence for this,” but noted that “these data collections did not show me what they [sic] students felt they had learned.”

When a plethora of quantitative data were collected, teachers often discovered these data were not as useful for answering the research questions as anticipated. For example, teachers collected a variety of quantitative data, including homework grades, in-class exam scores, survey responses, and standardized test scores. After collecting these data, they then had to identify which data best answered their research questions. In such instances, teachers reported results from a subset of the data. Eugene reflected that in the beginning he was “not sure as to how much data...[was] needed” and had “decided it was better to collect more data than...needed rather than not have enough.”

Teachers’ use of statistics to incorporate randomization For their action research projects, teachers almost always interviewed a subset of students. There are several ways to collect such samples, including purposeful and random sampling techniques (Mills, 2010), and the teachers were advised by their *Teacher as Scholarly Practitioner* instructors to use a method of interviewee selection that aligned with the research question(s) asked. For example, if teachers were making a change to support a particular subgroup of students, they were told to interview students from the targeted subgroup. Teachers studying cooperative learning strategies were encouraged to interview groups of students, selecting at least one group that was representative of the class and at most one that represented the variation in the class but was not typical of how other groups were behaving. Alternatively, if teachers were making a general change to their classroom practices, they were advised to divide their class into two to three groups based on mathematical achievement (or other variable of interest) and then randomly select students from each group to be interviewed.

In the action research projects reviewed, several teachers ($n = 6$) reported that they randomly selected students to interview. One of these teachers, Eliza, described how this was accomplished by obtaining “...pre and post interviews...of 6 random students whose names were drawn from a hat.” Others reported interviewing “students in a random order” (Spencer) or interviewing “random” (Eugene) or “randomly selected” (Melia) students with little to no additional explanation of the sampling process employed.

Teachers’ statistical thinking when planning studies and collecting data By collecting quantitative data, teachers exhibited a recognition of the need for such data. In all but one of the action research projects reviewed, quantitative data played either a complementary or primary role. However, the quantitative data were not always deliberately collected with an intended purpose in mind, leaving

teachers, such as Eugene, with additional data that may not have answered their intended research questions. Their awareness and desire to incorporate quantitative data suggests a beginning appreciation for that type of information, but leaves room for developing skills needed when planning a study to identify which type of data would help them address their questions of interest.

Teachers also exhibited an awareness of the presence of variation when they chose to collect baseline data in addition to poststudy data in order to help control for student-level variability. Jessie explicitly noted the importance of collecting such data before students had exposure to the new teaching practice, but other teachers, such as Cora, missed opportunities to measure student- and/or classroom-level changes on some items over time by asking different questions on the pre- and poststudy instruments. Even though these teachers recognized the importance of collecting baseline data (as emphasized in the STAT-MLT course), mid-study changes limited the conclusions they could make and highlighted teachers' limited experiences planning studies.

Teachers' choices to incorporate randomization suggest a potential awareness of the need to control for variability, but the ways in which they used randomization were primarily discussed at a superficial level that ignored whether such randomization fitted the study's purpose and the information a teacher hoped to obtain. Eliza was the only teacher who chose to include explanations about what made her selection process random rather than haphazard. "Randomly" selecting students or interviewing students in a "random" order appeared to be used as a buzzword intended to give the research integrity without further explanation about how it was used and/or fit the broader goals and purpose for the research.

6.2. ANALYZING DATA

Teachers' use of statistics to determine levels of analysis When analyzing their quantitative data, teachers made many decisions about how to summarize and report their data. As evidenced across the action research projects reviewed, one decision involved the level of analysis used to summarize the data; some data were reported at the individual student level, whereas other data were reported at the classroom (or multiple classes) level. For instance, Gretta numerically summarized performance at the classroom level by noting, "At the beginning of the research, about 10% of the students were correctly answering these problems, but by April around 80% of the students were correctly answering the 'Error Analysis' problems." Yet, in another example, Eugene used a bar chart to display changes in individual student test scores over the course of the project (see Figure 3), as well as two pie charts to summarize changes in standardized test performance at the classroom level (see Figure 4).

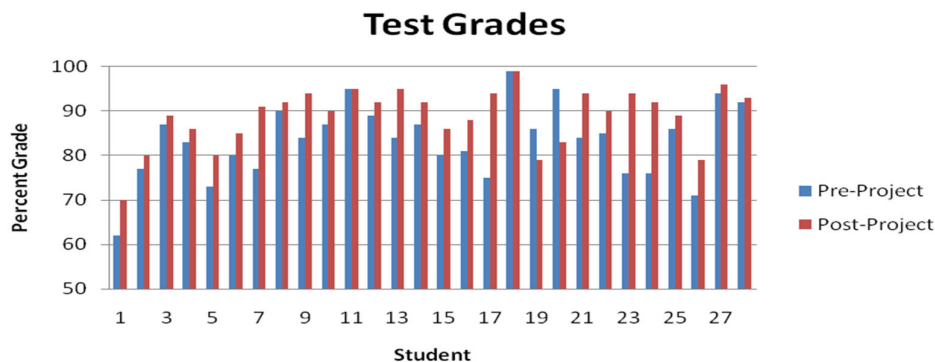
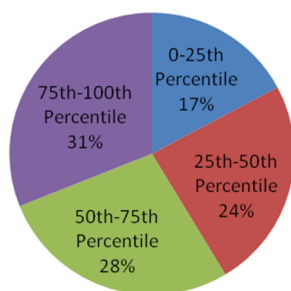


Figure 3. Eugene's graph of students' pre- to postproject test grades

Teachers' use of statistics to summarize data graphically and numerically In addition to choosing levels of analysis, teachers made decisions about which summaries to report and how to represent them. Student surveys were commonly used to collect data, and there were several instances in which survey items had ordinal response scales. In the papers reviewed, teachers made different choices about how to summarize ordinal data. As illustrated in Figure 5, some chose to analyze these data categorically, reporting frequencies for each response option, whereas others chose to summarize them quantitatively (see Figure 6), assigning numeric values to each response option.

Stars Math Pre-Project Test



Stars Math Post-Project Test

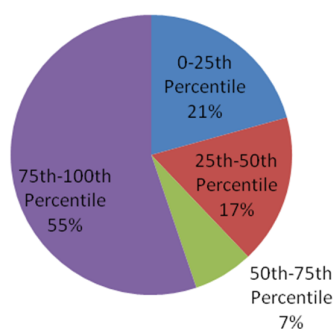


Figure 4. Eugene's graphs of students' pre- to postproject standardized test performance

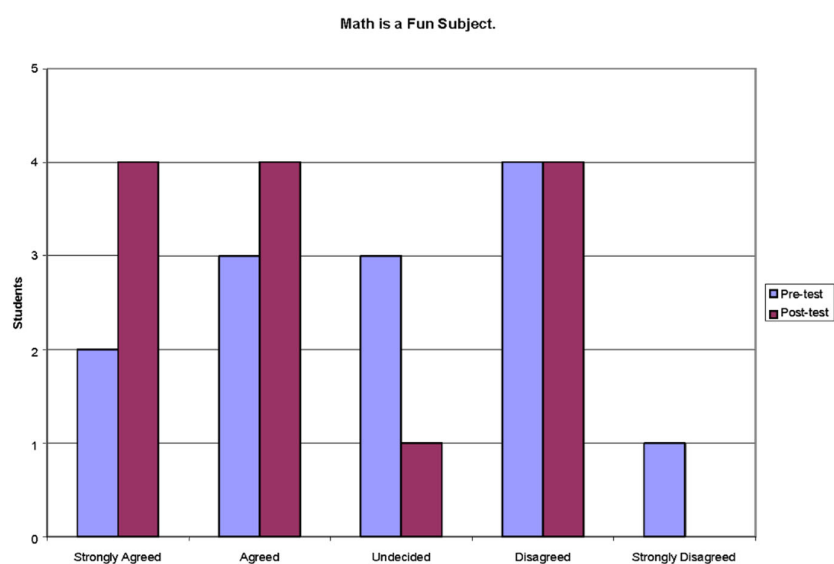


Figure 5. Maria's bar chart of pre- and postsurvey frequencies

	Strongly Agree 5	Agree 4	No Opinion 3	Disagree 2	Strongly Disagree 1	Pre Mean	Post Mean
I like to answer questions asked by the teacher in math class.						2.8	3.47
I feel comfortable asking questions in math if I don't understand something.						3.375	3
I am comfortable sharing my mathematical ideas with the whole class.						3.18	3.58
I am comfortable sharing my mathematical ideas in a small group.						3.8	4.23
I like to go to the board to present my math solutions to the whole class.						3.18	2.47

Figure 6. Eliza's pre- and postsurvey means table

Teachers also made different choices about how to format and present their information graphically and numerically. Commonly, teachers reported means, with some teachers choosing to also include standard deviations and others not. These numerical summaries were provided in text and/or in a visual display, such as a table (e.g., Figure 6) or a graph (e.g., Figure 7). Figure 6 represents a table that offers the survey items and scale, as well as the pre- and postsurvey mean responses for each item, but does not include standard deviations. Similarly, Figure 7 provides the mean responses for three different survey items, but with a graphical display instead of a table. In other instances, teachers only reported this information in the body of the text, conveying it with statements such as Jessie's: "The mean score in February was 3.55, in March it was 4.3, and in April it was 4.05."

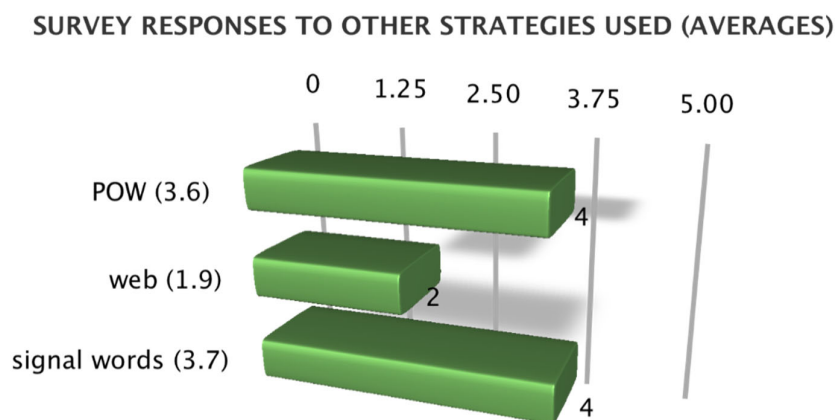


Figure 7. Eugene's default Microsoft Excel bar chart

When representing their data graphically, teachers primarily used pie charts (e.g., Figures 4 & 8), bar charts (e.g., Figures 3, 5 & 7), and line graphs (e.g., Figure 9). These graphs represent default graphing options in Excel, which teachers noted presented some challenges. For example, Eugene noted issues with the software, commenting that "not all of the graphs started with zero because of the program that created these graphs." The graphs in Figure 9 demonstrate these differences in the vertical axes. Eugene tried to creatively overcome this challenge by starting some of the graphs "with a grade of zero just to make sure that the graph started at zero like the others." He went on to say that the software "tried to eliminate unneeded graph space. If they all started with zero, it would be easier to compare them and see the increase in scores."

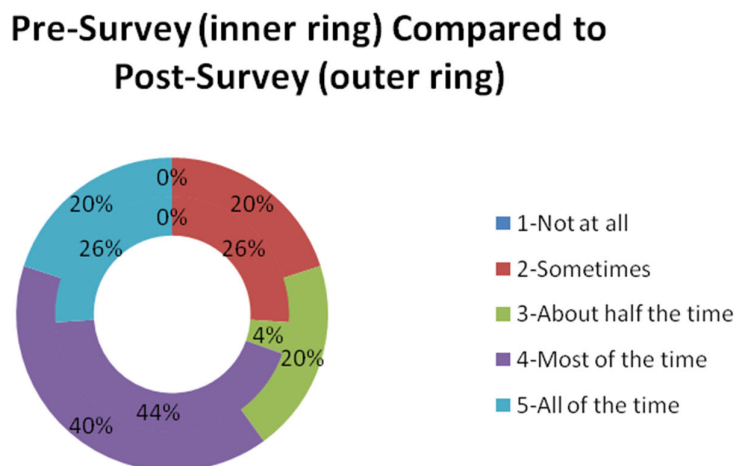


Figure 8. Melia's default Microsoft Excel pie chart

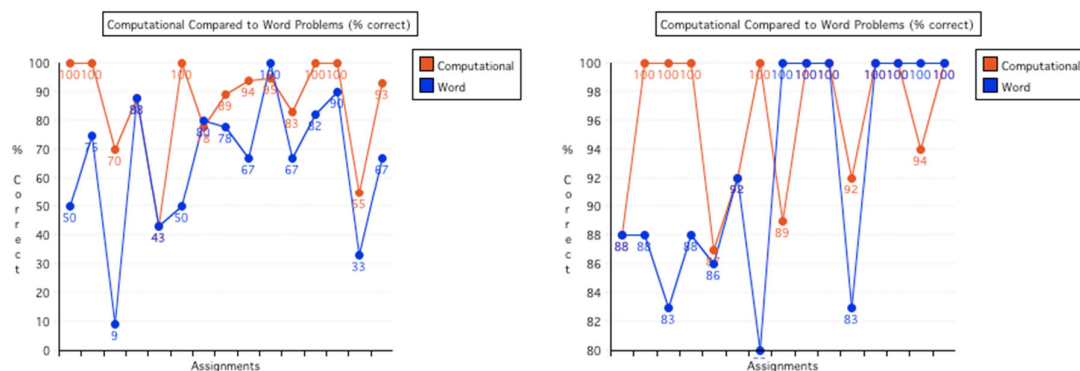


Figure 9. Example of Eugene's graphs with different scales

Teachers' statistical thinking when analyzing data When analyzing data, teachers chose different graphical "models" or representations to visualize their results at different levels (e.g., student vs. classroom level). The teachers often used graphs that were familiar to them (e.g., bar charts, pie charts, line graphs) and were easily produced in Microsoft Excel, even though the graphs may not have been the best displays to use or what was directly taught in STAT-MLT. The ways in which teachers chose to represent and summarize their data did not necessarily allow them to extract useful information. For example, by summarizing the data represented in Figure 4 at the classroom level, Eugene was unable to identify how individual students' scores on the STARS math test changed over time and whether or not the initial 31% of students who were in the 75th-100th percentile before the project began were still at that same percentile range after the project ended. In other cases, the graphs were difficult to read due to distortions of scale and/or other default plotting options (e.g., Figures 7 & 8), leading to potentially misleading results and/or comparisons. Further, some teachers, such as Eliza, chose to focus on means alone, ignoring additional information measures of variability would have provided (see Figure 6).

The figures provided in the action research projects suggest teachers did not necessarily explore different representations of the data in order to find which were the most informative. In one exception, Eugene described his process of representing the same data in different ways:

It was difficult to see any improvement by looking at just one assignment... Therefore, I tried using this table and looking at outcomes of all assignments together. For the most part, I still was not able to see much as far as any significant improvement by my students in the area of solving word problems. When I went back and found the differences between the average of computational problems they got correct and the average of word problems that they got correct, I was able to see some improvement.

He then went one step further and created line graphs to compare the class means for each type of problem across the different assignments. Through his explanations, Eugene reveals the dynamic process he used to explore different ways to summarize and visualize the same data in order to extract useful information about how his students' mean scores changed over time. This provides early evidence of transnumeration that was not evident in other teachers' action research projects. Overall, their numerical and graphical summaries often provided limited information, suggesting most teachers did not engage in the process of transnumeration.

6.3. INTERPRETING RESULTS

Teachers' use of statistics to determine scope of inference The action research projects reported statistics that were descriptive in nature, and these findings helped shape teachers' conclusions. In particular, through their interpretations and conclusions, teachers communicated to whom their results applied, that is, their scope of inference. In some instances, teachers made statements about their students alone. Eugene claimed, "My students' interest in math increased over time." Eugene noted finding "that some [reading strategies] worked better than others" and then continued to address class-

to-class variability by additionally noting that “this is something that may change from class to class.” In other instances, teachers made conclusions about students in general. For example, Cora asserted, “The data suggests that 65% of all students have increased confidence in their math abilities when working in teams.”

Teachers’ use of statistics to state potential limitations and alternative explanations When summarizing results and stating conclusions, teachers also included discussions of potential limitations and/or possible explanations for the observed results. For example, Eliza used students’ normal curve equivalencies on the mathematics subscale of the Terra Nova Achievement test at the end of two consecutive grades to detect changes in understanding as a result of the study. Although average gains were observed, Eliza acknowledged “it cannot be clearly determined from these test scores alone that increased discourse leads to better understanding.” She offered an alternative possible explanation, stating, “It could be that the curriculum is sufficiently covering the concepts on the test in order to show progress from year to year despite the method of delivery.”

In another action research project, Gretta wanted to gauge whether students could accurately assess their own understanding of mathematics. Before an assessment, each student recorded the score they predicted they would receive. The students’ predictions were then compared to their actual assessment scores. At the beginning of the study, there were discrepancies between students’ predictions and their performance. Gretta reported, “I gave students the first survey...and a majority...said that they would score a 3 out of 4 on the assessment...However,... most of the students scored a 4.” By the end of the study, most of the students accurately predicted their scores; for instance, Gretta wrote, “When I gave the survey again..., nearly all students said that they thought they would score a 4 on the assessment. It turned out that all 30 students scored a 4.” Gretta offered one explanation, concluding that “the students were able to do a much better job of assessing their understanding of the math.”

Teachers’ statistical thinking when interpreting results Teachers often synthesized their contextual understanding of their classrooms with the information the statistical results provided. For example, with his knowledge of how classes vary, Eugene recognized that his observed results “may change from class to class.” However, teachers did not consistently reference the data and acknowledge the role of variation when making conclusions, such as when Cora extended her results to all students.

Teachers also acknowledged constraints and limitations to varying degrees, with some exhibiting dispositions of skepticism, a propensity to seek deeper meaning, and logicalness. Some teachers, such as Eliza, questioned their results, proposing alternative explanations for the observed results, admitting limitations and logically reasoning through what conclusions could be drawn from the research results. Others ignored variability, such as when Cora generalized her results to all students, and/or did not explore alternative explanations for the observed results, such as when Gretta ignored the possibility that students learned to expect they would score a 4 on the assessment. Even though the teachers had similar experiences in STAT-MLT, these instances suggest teachers’ development of statistical dispositions may not be entirely course dependent and vary along a spectrum, ranging from naïve to more sophisticated levels.

7. DISCUSSION

These results begin to characterize the ways in which in-service middle-level teachers used statistics to conduct action research. The teachers’ varying uses of statistics were initially categorized in alignment with the phases of the statistical problem solving process: 1) planning studies and collecting data, 2) analyzing data, and 3) interpreting results. Themes identified within those categories and corresponding examples illustrate the specific and differing ways in which the teachers used statistics to conduct classroom research. These results help us to better understand how teachers use statistics to study and communicate the impact of changes in their classroom practice, as well as teachers’ development in their statistical thinking.

Our results suggest that teachers’ required knowledge of statistics is multi-faceted, requiring both a pedagogical component and statistical knowledge for the teaching profession. Teachers need to understand the statistics they teach so that they can design and provide meaningful learning opportunities for their students. However, as professionals, they also need to understand and use

statistics in meaningful ways that allow them to make informed decisions about their curriculum, instruction, and teaching practices. In short, the statistical content knowledge required to teach grade-level content standards for statistics differs from that required for using data to make decisions about classroom practices. The varying ways in which the teachers implemented statistical concepts during action research required more nuanced uses of their statistical knowledge. For example, in addition to knowing how to teach middle-level students appropriate ways to graphically and numerically summarize data, the mathematics teachers needed to know how to present information in ways that allowed them to meaningfully answer their research questions and, ultimately, make informed decisions about their classroom practices. Consequently, teachers' specialized content knowledge for teaching statistics (Groth, 2013) may encompass both content knowledge that directly supports classroom teaching and knowledge that supports classroom-based research, or data-driven decision making, for improving instruction and student learning.

The ways in which teachers used statistics during their action research projects also provided evidence of teachers' early development in statistical thinking. Using Wild and Pfannkuch's (1999) model of statistical thinking "as a tool for evaluating student thinking; and as a reference point against which to check learning opportunities provided to students" (p. 246), we explored teachers' statistical thinking across the four-dimensional framework. Because of the nature of research, teachers engaged in the entire statistical problem solving process throughout their action research projects. However, teachers demonstrated different types of thinking fundamental to statistical thinking. The action research projects revealed that teachers consistently recognized the need for data and interpreted it within the context of their classrooms, but they differed in their consideration of variability, with some ignoring it altogether or referencing it superficially through "randomization" and others explicitly discussing and accounting for its role in their results. In their projects, teachers used graphical and numerical summaries to represent their data, but these summaries were often familiar depictions that were easy to obtain rather than meaningful summaries that generated a deeper understanding of the information the data provided. In the written projects, we saw little to no evidence of teachers engaging in the process of transnumeration and the interrogative cycle. Teachers also showed differing levels of skepticism and logic, but most dispositions were not evident based on what they wrote. In general, teachers demonstrated beginning levels of statistical thinking, highlighting a need for educational opportunities that better support teachers' development of statistical thinking in their multi-faceted work as both instructors and users of statistics. Together, our results have important implications for statistics educators, teacher educators, and statistics education researchers.

7.1. IMPLICATIONS FOR STATISTICS EDUCATORS

These findings have important implications for how we, as statistics educators, can support middle-level mathematics teachers' continued growth and development in statistics. We designed the STAT-MLT course to focus on aspects of statistical thinking throughout the statistical problem solving process that we thought teachers would need as both educators and users of statistics, and we hoped that would be translated to their action research projects. After reviewing teachers' action research projects, we learned we had not adequately prepared teachers to use statistics and engage in the statistical problem solving process to study their problems of practice. Although teachers had conducted a data analysis project during the STAT-MLT course, they did not have the opportunity to engage in all phases of the statistical problem solving process; in the course, teachers were provided a data set, so they did not engage in the phase of planning a study and collecting data. Furthermore, as evidenced by the action research projects, teachers were interested in making decisions about their particular students and/or classes, and descriptive statistics were most appropriate to use in those situations.

Our results suggest statistics educators need to continue to immerse and support teachers in the early phases of planning studies and collecting data if they are to be successful in implementing these methods in their classrooms (CBMS, 2012; U.S. Department of Education, Office of Planning, Evaluation and Policy Development, 2011). We should structure learning opportunities that allow teachers to engage in data-driven decision making, where they must collect, summarize, and use student- and classroom-level data to make informed decisions about their classroom practices. As also recommended by Pfannkuch and Ben-Zvi (2011), such learning opportunities should include targeted discussions about designing studies, identifying and collecting data for answering particular questions

of interest, and determining appropriate levels of data analysis (e.g., student- and/or classroom-level) and scope of inference. In addition, our study shows that more time should be devoted to the practice of summarizing data graphically and numerically, immersing teachers in the process of transnumeration in which they must create, explore, and critically evaluate multiple representations in order to uncover a deeper understanding of the data and the information they provide. Even further, courses for teachers could also address small-sample techniques, rather than focusing solely on formal large-sample inference procedures. Although the goal of this study was not to evaluate the learning outcomes of the course, we gained valuable insight into the ways teachers used statistics and what their uses revealed about their statistical thinking. Future research should explore the impact course content and structure have on teachers' development of statistical thinking when conducting classroom-based research.

7.2. IMPLICATIONS FOR TEACHER EDUCATORS

These findings also have important implications for teacher educators. Our results highlight the need to create more authentic learning opportunities in statistics for teachers (Makar & Confrey, 2005), some of which may extend beyond traditional coursework. Action research projects, or other classroom-based research experiences, provide experiential learning opportunities that, with support and mentorship from statistics educators, could allow teachers to continue developing their statistical thinking. When teachers choose to use quantitative data to answer questions about their classroom practices, they are immersed in the statistical problem solving process and data-driven decision making. With statistical thinking playing a foundational role throughout this entire process, action research provides rich opportunities to support and explore teachers' statistical development. Statistics educators should be involved throughout the whole investigative process, potentially serving on a team with mathematicians and/or mathematics educators, as recommended by Franklin et al. (2015). This team of advisors can then help guide teachers through the entire research experience, providing additional mentorship and direction in a variety of research methodologies, including statistics.

7.3. IMPLICATIONS FOR STATISTICS EDUCATION RESEARCHERS

Teachers' action research projects provide us additional insight into teachers' statistical thinking and development within the context of their own classroom inquiry, and we posit a valuable means to better understand the nuanced complexities of teachers' required statistical knowledge for teaching. Mathematical knowledge for teaching (MKT) was motivated by a need to understand the mathematics/mathematical knowledge that teachers actually use in classrooms. Ball and colleagues have referred to this as a "practice-based" understanding of mathematics (e.g., Ball & Bass, 2003; Ball, Thames, & Phelps, 2008). It is knowledge that is mathematical in nature and is a part of pedagogy (as in, the kinds of topics that would be covered in a traditional methods class). Some have gone so far as to call it as a form of applied mathematics (Stylianides & Stylianides, 2010). We, along with Groth (2013) are seeking a similar conception of statistical knowledge—that is "statistical knowledge for teaching." In some ways, this notion overlaps with MKT in that it involves knowledge teachers use while teaching statistics. A key difference, however, is that teachers also use statistics to analyze and understand their practice and to make decisions about a myriad of topics ranging from student placement to instructional practices to curricular decisions (Jacobs, Gregory, Hoppey, & Yendol-Hoppey, 2013). This statistical knowledge for teaching involves learning to use statistics wisely, carefully, and ethically as a part the continuous cycles of improvement in which teachers are constantly engaging.

7.4. LIMITATIONS AND FUTURE RESEARCH

In this study, statistics educators were not actively involved in the mentoring and action research process. Statisticians studied the action research projects after they were completed rather than while teachers were immersed in the research process, so the findings are limited to what teachers chose to write in their reports. Consequently, we chose not to investigate teachers' research questions or evaluate whether or not their uses of statistics were appropriate for directly answering those questions; after the first cohort, the mathematics educators were much more directive with teachers' research questions and

data collection efforts, so such an analysis was not comparable across cohorts. In future research, statistics education researchers should be involved during the entire investigative process to capture more broadly teachers' statistical thinking and development throughout the duration of the action research project.

8. CONCLUSION

In this paper, we investigated how in-service middle-level mathematics teachers used statistics to study the impact of changes in classroom practice on their teaching and their students, and what those uses revealed about teachers' development of statistical thinking. The corresponding results suggest that teachers' required knowledge of statistics extends beyond teaching, providing insight into how in-service teachers use statistical concepts to assess students and their own practice. Such findings help us better understand how teachers use statistics to make classroom-based decisions informed by data, enabling us to structure learning opportunities that help meet teachers' professional needs. Future research should expand on Groth's (2013) theoretical framework for and development of mathematics teachers' statistical knowledge for teaching, addressing their statistical needs as teaching professionals as well as their pedagogical needs.

ACKNOWLEDGEMENTS

This work was supported in part by a grant from the National Science Foundation, DUE-0412502. The findings and opinions are those of the authors and not necessarily those of the funding agency.

REFERENCES

- Advisory Committee on Mathematics Education (ACME). (2002). *Continuing professional development for teachers of mathematics*. London, England: Author.
[Online: <http://www.ncetm.org.uk/public/files/7072161/Professional-Development-Opportunities-For-Mathematics-Teachers-in-Post16.pdf>]
- Ball, D. L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 Annual Meeting of the Canadian Mathematics Education Study Group*, (pp. 3–14). Edmonton, Canada: CMESG/GCEDM.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407.
- Bargagliotti, A. E. (2014, January). The assessment of teachers: Notes from a conference. *Amstat News*, 439, 30–31.
- Batanero, C., Burrill, G., & Reading, C. (2011). Overview: Challenges for teaching statistics in school mathematics and preparing mathematics teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education: A Joint ISMI/IASE Study* (pp. 407–418). New York: Springer.
[Online: <http://www.ugr.es/~batanero/pages/ARTICULOS/Overview.pdf>]
- Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 3–15). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: Analysis of a complex cognitive skill. *Computers in Human Behavior*, 21, 487–508.
[Online: <https://www.ou.nl/Docs/Expertise/OTEC/Publicaties/iwan%20wopereis/CIHB-artikel%20Brand-Gruwel-Wopereis-Vermetten.pdf>]
- Chance, B. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education*, 10(3).
[Online: <http://www2.amstat.org/publications/jse/v10n3/chance.html>]
- Chen, S., Huang, F., & Zeng, W. (2017). Comments on systematic methodologies of action research in the new millennium: A review of publications 2000–2014. *Action research*.
[Online: <http://journals.sagepub.com/doi/pdf/10.1177/1476750317691103>]

- Cochran-Smith, M., & Lytle, S. L. (1993). *Inside/outside: Teacher research and knowledge*. New York: Teachers College Press.
- Cochran-Smith, M., & Lytle, S. L. (2009). *Inquiry as a stance: Practitioner research for the next generation*. New York: Teachers College Press.
- Common Core State Standards Initiative. (2010). *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.
[Online: <http://www.corestandards.org>]
- Conference Board of the Mathematical Sciences. (2001). *The mathematical education of teachers*. Providence, RI and Washington, DC: American Mathematical Society and Mathematical Association of America.
- Conference Board of the Mathematical Sciences. (2012). *The mathematical education of teachers II*. Providence, RI and Washington, DC: American Mathematical Society and Mathematical Association of America.
- Franklin, C. (2013). Common Core Standards and the future of teacher preparation in statistics. *The Mathematics Educator*, 22(2), 3–10.
- Franklin, C. A., Bargagliotti, A. E., Case, C. A., Kader, G. D., Scheaffer, R. L., & Spangler, D. A. (2015). *The statistical education of teachers*. Arlington, VA: American Statistical Association.
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework*. Alexandria, VA: American Statistical Association.
[Online: https://www.amstat.org/asa/files/pdfs/GAISE/GAISEPreK-12_Full.pdf]
- Froelich, A., Kliemann, W., & Thompson, H. (2008). Changing the statistics curriculum for future and current high school mathematics teachers: A case study. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*, Monterrey, Mexico.
[Online: https://iase-web.org/documents/papers/rt2008/T6P5_Froelich.pdf]
- Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 47–78). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Garfield, J. (1995). How students learn statistics. *International Statistical Review*, 63(1). 25–34.
[Online: <http://www.jstor.org/stable/1403775>]
- Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, 10(3).
[Online: <http://ww2.amstat.org/publications/jse/v10n3/garfield.html>]
- Garfield, J., & Ben-Zvi, D. (2009). Helping students develop statistical reasoning: Implementing a statistical reasoning learning environment. *Teaching Statistics*, 31(3), 72–77.
[Online: <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9639.2009.00363.x/abstract>]
- Garfield, J., & delMas, R. (2010). A web site that provides resources for assessing students' statistical literacy, reasoning, and thinking. *Teaching Statistics*, 32(1).
[Online: onlinelibrary.wiley.com/doi/10.1111/j.1467-9639.2009.00373.x/abstract]
- Groth, R. E. (2013). Characterizing key developmental understandings and pedagogically powerful ideas within a statistical knowledge for teaching framework. *Mathematical Thinking and Learning*, 15(2), 121–145.
- Hamilton, L., Halverson, R., Jackson, S. S., Mandinach, E., Supovitz, J. A., & Wayman, J. C. (2009). Using student achievement data to support instructional decision making (NCEE 2009-4067). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
[Online: https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/dddm_pg_092909.pdf]
- Heaton, R. M., Lewis, W. J., Homp, M. R., Dunbar, S. R., & Smith, W. M. (2013). Challenging yet accessible mathematics courses for middle level teachers. In C. Beaver, L. Burton, M. Fung, & K. Kruczek (Eds.), *Resources for Preparing Middle School Mathematics Teachers*, Mathematical Association of America Note Series, (pp. 181–201). Washington, DC: MAA Press.

- Heaton, R. M., Lewis, W. J., & Smith, W. M. (2009). Building middle level mathematics teachers' capacities as teachers and leaders: The Math in the Middle Institute Partnership. *The Journal of Mathematics and Science: Collaborative Explorations*, 11, 1–18.
- Heaton, R. M., Lewis, W. J., & Smith, W. M. (2013). The Math in the Middle Institute: Strengthening middle level teachers' mathematical and pedagogical capacities. In C. Beaver, L. Burton, M. Fung, & K. Kruczek (Eds.), *Resources for Preparing Middle School Mathematics Teachers*, Mathematical Association of America Note Series, (pp. 47–58). Washington, DC: MAA Press.
- Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the “data-driven” mantra: Different conceptions of data-driven decision making. *Yearbook of the National Society for the Study of Education*, 106, 105–131.
[Online: <https://www.rand.org/pubs/reprints/RP1372.readonline.html>]
- Jacobs, J., Gregory, A., Hoppey, D., & Yendol-Hoppey, D. (2013). Data literacy: Understanding teachers' data use in a context of accountability and response to intervention. *Action in Teacher Education*, 31(3), 41–55.
- Kazak, S., & Confrey, J. (July, 2004). *Investigating education practitioners' statistical reasoning in analysis of student outcome data*. Paper presented at the Tenth International Conference of Mathematical Education, Copenhagen, Denmark.
[Online: <https://iase-web.org/documents/papers/icme10/KazakConfrey.pdf>]
- Lachat, M. A., & Smith, S. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed At Risk*, 10(3), 333–349.
- Lai, M. & Robinson, V. (2006). *Practitioner research for educators: A guide to improving classrooms and schools*. Thousand Oaks, CA: Corwin.
- Levitt, S. D., & Dubner, S. J. (2009). *Freakonomics: A rogue economist explores the hidden side of everything*. New York, NY: HarperCollins.
- Makar, K. & Confrey, J. (2004). Secondary teachers' statistical reasoning in comparing two groups. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 353–373). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Makar, K. & Confrey, J. (2005). “Variation-Talk”: Articulating meaning in statistics. *Statistics Education Research Journal*, 4(1), 27–54.
[Online: [https://iase-web.org/documents/SERJ/SERJ4\(1\)_Makar_Confrey.pdf](https://iase-web.org/documents/SERJ/SERJ4(1)_Makar_Confrey.pdf)]
- Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist*, 47(2), 71–85.
- Mandinach, E. B., & Gummer, E. S. (2012). *Navigating the landscape of data literacy: It IS complex*. Washington, DC & Portland, OR: WestEd and Education Northwest.
- Marriott, J., Davies, N., & Gibson, L. (2009). Teaching, learning and assessing statistical problem solving. *Journal of Statistics Education*, 17(1).
[Online: <https://www2.amstat.org/publications/jse/v17n1/marriott.pdf>]
- Marsh, J. A. (2012). Interventions promoting educators' use of data: Research insights and gaps. *Teachers College Record*, 114(11), 1–48.
- Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). Making sense of data-driven decision making in education: Evidence from recent RAND research. *Rand Occasional Papers*, 1-16. Santa Monica, CA: RAND corporation.
[Online: https://www.rand.org/pubs/occasional_papers/OP170.html]
- Means, B., Chen, E., DeBarger, A., & Padilla, C. (2011). *Teachers' ability to use data to inform instruction: Challenges and supports*. Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development.
- Miller, N. (2017). Teaching inquiry to high school teachers through the use of mathematics action research projects. *PRIMUS*, 27(1), 33–46.
[Online: <http://dx.doi.org/10.1080/10511970.2016.1199616>]
- Mills, G. E. (2010). *Action research: A guide for the teacher researcher* (4th Ed.). Upper Saddle River, NJ: Prentice Hall.
- National Centre for Excellence in the Teaching of Mathematics (NCETM). (2011). *Professional development opportunities for mathematics teachers in post 16 colleges*. Sheffield, UK: Author.
[Online: <https://www.ncetm.org.uk/public/files/7072161/Professional-Development-Opportunities-For-Mathematics-Teachers-in-Post16.pdf>]

- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- Pfannkuch, M., & Ben-Zvi, D. (2011). Developing teachers' statistical thinking. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education: A Joint ISMI/IASE Study*, (pp. 323–333). New York: Springer.
- Rossmann, A., Medina, E., & Chance, B. (2006). A post-calculus introduction to statistics for future secondary teachers. In A. Rossmann & B. Chance (Eds.), *Proceedings of the 7th International Conference on Teaching Statistics (ICOTS7)*, Voorburg, The Netherlands: International Statistical Institute.
- Rumsey, D. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education*, 10(3).
[Online: <http://ww2.amstat.org/publications/jse/v10n3/rumsey2.html>]
- Scheaffer, R. L., & Jacobbe, T. (2014). Statistics education in the K-12 schools of the United States: A brief history. *Journal of Statistics Education*, 22(2).
[Online: <http://ww2.amstat.org/publications/jse/v22n2/scheaffer.pdf>]
- Schmid, K. K., Blankenship, E. E., Kerby, A. T., Green, J. L., & Smith, W. M. (2014). The development and evolution of an introductory statistics course for in-service middle-level mathematics teachers. *Journal of Statistics Education*, 22(3).
[Online: <http://ww2.amstat.org/publications/jse/v22n3/schmid.pdf>]
- Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester, Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 957–1009). Greenwich, CT: Information Age.
- Smith, W. M., & Heaton, R. M. (2013). Learning from practice about improving the quality of mathematics teacher research. *Mathematics Teacher Educator*, 1(2), 148–161.
- Stuart, M. (1995). Changing the teaching of statistics. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 44(1), 45–54.
[Online: <http://www.jstor.org/stable/2348615>]
- Stylianides, G. J., & Stylianides, A. J. (2010). Mathematics for teaching: A form of applied mathematics. *Teaching and Teacher Education*, 26(2), 161–172.
- University of Wisconsin-Madison School of Education. (2017, February 28). *K-8 mathematics professional development opportunities*.
[Online: cop.education.wisc.edu/cop/professional-learning/courses/mathematics-microcredential]
- U.S. Department of Education, Office of Planning, Evaluation and Policy Development (2011). *Teachers' ability to use data to inform instruction: Challenges and supports*. Washington, DC: Author.
[Online: <http://www.ed.gov/about/offices/list/opepd/ppss/reports.html>]
- Vaughan, M., & Burnaford, G. (2016). Action research in graduate teacher education: A review of the literature 2000–2015. *Educational Action Research*, 24(2), 280–299.
- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–265.
[Online: <https://iase-web.org/documents/intstatreview/99.Wild.Pfannkuch.pdf>]
- Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.
- Zeichner, K. (2001). Educational action research. In P. Reason & H. Bradbury (Eds.), *Handbook of action research: Participative inquiry and practice* (pp. 273–283). Thousand Oaks, CA: Sage.

JENNIFER L. GREEN
Department of Mathematical Sciences
Montana State University
2-214 Wilson Hall
P.O. Box 172400
Bozeman, MT 59717
USA