The Development and Evolution of an Introductory Statistics Course for In-Service Middle-Level Mathematics Teachers

Kendra K. Schmid
University of Nebraska Medical Center

Erin E. Blankenship
University of Nebraska Lincoln

April T. Kerby
Winona State University

Jennifer L. Green
Montana State University

Wendy M. Smith
University of Nebraska Lincoln

Key Words: Statistics education research; Middle-level mathematics in-service.

Abstract

The statistical preparation of in-service teachers, particularly middle school teachers, has been an area of concern for several years. This paper discusses the creation and delivery of an introductory statistics course as part of a master's degree program for in-service mathematics teachers. The initial course development took place before the advent of the Common Core State Standards for Mathematics (CCSSM) and the Mathematics Education of Teachers (MET II) Reports, and even before the GAISE Pre-K-12 Report. Since then, even with the recommendations of MET II and the wide-spread implementation of the CCSSM, the guidance available to faculty wishing to develop a statistics course for professional development of in-service teachers remains scarce. We give an overview of the master's degree program and

discuss aspects of the course, including the goals for the course, course planning and development, the instructional team, course delivery and modifications, and lessons learned through five offerings. With this paper, we share our experiences developing such a course, the evolution of the course over multiple iterations, and what we have learned about its value to the middle-level teachers who have participated. As more and more universities are being asked to develop courses specifically for in-service teachers, we wrote this article in the hopes of providing guidance to others, and to share our lessons learned.

1. Introduction

This paper discusses the creation and delivery of an introductory statistics course as part of a master's degree program for in-service mathematics teachers. We give an overview of the program and discuss aspects of the course, including course planning and development, the goals for the course, the instructional team, course delivery and modifications, and lessons learned through five offerings.

1.1 Literature

The statistical preparation of in-service teachers, particularly middle school teachers, has been an area of concern for several years. The Mathematical Education of Teachers (MET) Report (Conference Board of the Mathematical Sciences - CBMS 2001) asserts that "Of all the mathematical topics now appearing in middle grades curricula, teachers are least prepared to teach statistics and probability" (p. 114). While the focus of the statistics courses recommended by the MET II Report (CBMS 2012) has evolved since the first report (CBMS 2012, p. 18), the importance of a sound statistical education for teachers has not. The Common Core State Standards (CCSS) were released in 2010 and have been adopted by many states (CCSS 2010). The standards for mathematics, CCSSM, place greater emphasis on statistics and probability than have previous standards documents, and most of the CCSSM statistical content is introduced in the Grades 6-8 standards. The MET II Report (CBMS 2012) recognizes this shift in content placement, and states "Many teachers prepared before the era of the CCSS will need opportunities to study content that they have not previously taught, particularly in the areas of statistics and probability" (p. 68).

Our initial course development took place before the advent of the CCSSM and MET II, and even before the GAISE Pre-K-12 Report (Franklin et al. 2007). At the time, little had been published regarding the statistical education of teachers, particularly in-service teachers. Most of what had been published seems, in retrospect, antiquated. The MET Report suggests that "…teachers must develop both skills for calculation and those for interpretation" (p. 114) and describes much of the middle school statistics curriculum in the context of probability (CBMS 2001). This focus on probability was not isolated to the MET Report. The literature that could be found at that time on both teacher training (Batanero, Godino, and Roa 2004) and the use of statistics in the middle school classroom (e.g., Abrahamson, Janusz, and Wilensky 2006) also centered on probability.

Despite this emphasis on probability for teacher training, other topics in statistics, such as data analysis, were gaining attention in K-12 mathematics standards. Beginning with the early efforts of the ASA-NCTM Quantitative Literacy Project (Scheaffer 1986; 1991), the National Council

of Teachers of Mathematics (NCTM) introduced statistics as one of the five mathematics content strands (NCTM 1989; Scheaffer and Jacobbe 2014). In 2000, NCTM released their new set of guidelines, *Principles and Standards for School Mathematics,* that recommend various standards related to data analysis and probability be taught throughout the K-12 mathematics curriculum. Extending beyond probability, the standards within this content area address other statistical concepts: "[1] **Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them…[2] Select and use appropriate statistical methods to analyze data**…[and 3] D**evelop and evaluate inferences and predictions that are based on data**" (NCTM 2000).

Around this same time, leaders within the statistics education community were calling for reforms in teaching undergraduate statistics. Cobb's (1992) recommendations to emphasize statistical thinking, use real data and implement active learning, sparked a new era of reform, particularly within the non-calculus-based introductory statistics course. In 1997, Moore described what should be changed in the statistics classroom including changing content to allow more data analysis and less probability, allowing more active learning, and the use of technology for data analysis and simulations. Even though these conversations tended to focus on introductory statistics courses for undergraduate students, these ideas and the reform movement were gaining momentum at the time we were developing our course (Garfield 2000, 2002), and they, in conjunction with the NCTM guidelines (2000), played a substantial role in how we designed and taught our course for middle-level teachers.

Post-GAISE, there has been some additional work on modernizing teacher training, but it tends to be focused on either pre-service teachers (Garfield and Everson 2009; Leavy 2010; Metz 2010) or evaluating a specific pedagogical technique (Lesser and Kephart 2011; Smith and Hialmarson 2013). Other work is available on middle school students and their understanding of specific concepts (Carmichael, Callingham, Watson, and Hay 2009; Lee, Angotti, and Tarr 2010; Watson 2008). Additionally, more has been published on teacher development programs in mathematics (e.g., Adler and Davis 2006; Schifter 1993; Schifter and Fosnot 1993; Simon 1994; Simon and Schifter 1991; Simon and Schifter 1993), but not focused on statistics. Statistics is a discipline separate from mathematics (Cobb 1993; Rossman, Chance, and Medina 2006); mathematicians and teacher educators providing professional development courses or workshops to teachers rarely have significant statistical backgrounds, and thus most are ill-poised to help teachers develop statistical thinking skills or even master the statistics content in their curricula. Not only is the statistical content in the middle-school curricula growing, but the pervasiveness of data in society is expanding. Including statistics as only one topic of a mathematics professional development course or workshop contributes to it remaining only a "special topic" in mathematics courses, leaving students unexposed to the role statistics play in everyday life. Through a course designed for middle school teachers that focuses solely on statistics, we as statisticians and statistics educators can demonstrate how statistical concepts can be reinforced in a variety of applications, such as science and social studies, and not just as a "special topic" in mathematics.

When we were asked to develop a course for in-service middle school teachers, we could not find an article, or set of articles, to guide us through the course development process. Over the past eight years, we have referenced relevant standards, guidelines and recommendations (e.g.,

MET II Report (CBMS 2012), CCSSM (2010), GAISE College (Aliaga et al. 2005) and PreK-12 (Franklin et al. 2007) Reports, Nebraska State Mathematics Standards http://nde.ne.gov/math/index.html) as they have been published and used them to help modify and improve our course. However, we still have not found references describing how to best structure/develop a statistics course for in-service middle school teachers. As more and more universities are being asked to develop courses specifically for in-service teachers, we wrote this article in the hopes of providing guidance to others, and to share our lessons learned.

1.2 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 (EHR-0142502). Participants completing the 36 credit hour program earn master's degrees through the Department of Mathematics (Master of Arts for Teachers [MAT]) or the Department of Teaching, Learning, and Teacher Education (Master of Arts [MA]). The goals of the Math in the Middle Institute, as found on the Institute website (http://scimath.unl.edu/MIM/index.php) are:

"Through their experience in Math in the Middle, teachers:

- built strong mathematics content knowledge
- studied and practiced the art of pedagogy
- developed the ability to conduct action research about their teaching practices
- cultivated leadership skills
- applied this knowledge and skill in their classrooms, schools and districts"

The M^2 Institute courses are a combination of on-campus summer courses and academic year distance courses. Due to differences in the timing of funding, cohorts of teachers have completed the 12 M^2 Institute courses over a period of 18 to 25 months; six cohorts of teachers participated in the M^2 Institute 2004 to 2011. While new funding received in 2013 has resulted in two additional Math in the Middle cohorts, this paper focuses on the development of the course and the first five cohorts.

The M^2 Institute includes seven mathematics courses, one statistics course, three education/pedagogy courses, and one capstone course. See Table 1 for a list of course titles; more information about the M^2 Institute courses, including the statistics course we developed, can be found on the M^2 Institute website (http://scimath.unl.edu/MIM/) and in previous publications about the M^2 Institute (Heaton, Lewis and Smith 2009, 2013; Heaton, Lewis, Homp, Dunbar and Smith 2013). Most cohorts of teachers have had 30-35 participants. Although the middle level focus of the M^2 Institute is on grades 5-8, due to the preponderance of rural school districts in Nebraska, many teachers teach math to multiple grades of students (e.g., K-6, 3-6, 7-8, 7-12). Between 2004 and 2011, 160 teachers from across the state of Nebraska participated in the M^2 Institute; these 160 teachers represent over 100 schools from more than 60 districts. 157 teachers earned master's degrees through the M^2 Institute from 2005 to 2011; this represents a 98% retention rate to graduation.

Table 1. Math in the Middle Institute courses

Course Number	Course Title
MATH 800T	Mathematics as a Second Language
MATH 802T	Functions, Algebra, and Geometry for Middle Level Teachers
TEAC 800	Inquiry into Teaching and Learning
MATH 804T	Experimentation, Conjecture and Reasoning
TEAC 801	Curriculum Inquiry
MATH 805T	Discrete Mathematics for Middle Level Teachers
MATH 806T	Number Theory and Cryptology for Middle Level Teachers
STAT 892	Statistics for Middle Level Teachers
TEAC 888	Teacher as Scholarly Practitioner
MATH 807T	Using Math to Understand Our World
MATH 808T	Concepts of Calculus
TEAC 889 or MATH 896[a]	Integrating the Teaching and Learning of Math (Capstone Course)

[a] Teachers pursuing the MAT degree register for MATH 896; teachers pursing the MA degree register for TEAC 889.

1.3 Purpose

As discussed in the literature review, the statistical/mathematical education of pre- and in-service teachers has been identified as a university, state and national priority. When our university received an NSF grant to develop the master's degree program, there was very little literature on how to best equip in-service teachers for statistical thinking, understanding educational data, or mastering the new statistical content they were being asked to teach. Since then, even with the recommendations of MET II (CBMS 2012) and the wide-spread implementation of the CCSSM (2010), the guidance available to faculty wishing to develop a statistics course for professional development of in-service teachers remains scarce. With this paper, we share our experiences developing such a course, the evolution of the course over multiple iterations, and what we have learned about its value to the middle-level teachers who have participated.

2. Course Development

The first cohort of teachers entered the master's program in fall, 2004, and we were approached in mid-March, 2005, about developing a course that would be offered for the first time in July, 2005. The course was to take place over one week at a local high school, so advanced planning was essential. We wouldn't have the luxury of extensive mid-course corrections; even photocopies had to be made before the course ever started. As a result, we spent far more time outlining course content than we would in a typical semester-long course. In addition, we needed to translate the material found in a traditional introductory statistics course to make it relevant for middle-level teachers to apply in a variety of ways. In this section we discuss the goals we had for the course, the instructional team, and the thought-process we went through as we were outlining the course.

2.1 Course Goals

The goals for the course were three-fold. First, the overarching goal of the M^2 Institute is to help teachers better understand mathematical concepts and how to relay mathematical ideas to their students. Likewise, this was our first goal for the course, to help the teachers improve their statistical reasoning and thinking, as well as their ability to help their students do the same. Second, we wanted to help them understand the statistical information they are exposed to in their careers, from standardized test scores to educational research. We thought it important that teachers be familiar with and understand the data they receive on their own students and schools as well as educational research in which they may be interested. The third goal of the course was to help prepare the teachers for their action research projects. During the second year of the M^2 Institute program, teachers are required to do an action research project. For this project they craft research questions and identify variables to study in their own teaching practices. They then design a research project, and collect and analyze data. It was a goal of this course to provide teachers with the skills they would need to analyze quantitative data for these projects. While we covered some of the topics in the Nebraska mathematics standards, it was not a primary goal of the course to prepare the teachers for that specific content.

2.2 Instructional Team

The instructional team for our course was composed of five instructors: a statistics faculty member, three statistics graduate students, and a master teacher. With our varying backgrounds and experiences, we were all able to contribute to different aspects of the course development and delivery. The statistics faculty member had extensive background in teaching undergraduate and graduate level statistics courses, and was approached to develop this course because she, among the faculty in the Department of Statistics, had the most interest in teacher development. The statistics graduate students all had experience teaching introductory statistics at the undergraduate level, and were proven teachers. The master teacher during the first two iterations of the course was a local high school teacher who had many years of experience teaching high school level/AP statistics, and was selected based on recommendations from the M^2 Institute leadership team. In subsequent course offerings, the master teacher was a graduate of the M^2 Institute, and a local middle school mathematics teacher who was selected based on outstanding performance in the class. The role of the master teacher was critical because of the ability to help shape the course from the viewpoint of someone with experience with middle-level classroom practice. Except for the master teacher, the instructional team was the same for the first four iterations of the course. In the 5th iteration, the instructional team included three of the original team members, along with two new team members to keep the same composition of instructors described above. This instructional team makeup was similar across other M^2 Institute courses so that we could offer teachers more individualized guidance and diverse perspectives. Most cohorts had between 30-35 teachers, and the courses were offered in relatively short periods of time (1-2 weeks), so the size and diversity of our instructional team gave us the flexibility to offer ample support for the teachers.

2.3 Planning

Course development was initiated about four months prior to the first course offering. The faculty member was responsible for the preliminary planning, such as determining course content and order of topics. In transitioning from *what* should be taught to *how* the content should be taught, the course development evolved into a team effort. Regular meetings were held to plan how each day of the course would be structured. Additionally, we consulted with colleagues in education and instructors of other Math in the Middle courses. Because all members of the instructional team had experience in the introductory statistics classroom and were familiar with the move toward educational reforms in statistics (e.g., Cobb 1992; Moore 1997), we all had ideas about strategies for effective classroom practice. While the faculty member and graduate students had experience with introductory statistics, we had not previously worked with in-service teachers. The master teacher was invaluable in sharing not only her experiences with pre-secondary/secondary students, but also the unique challenges of working with teachers. For example, based on her advice, we wanted to make sure that all course content could be connected in some way to the classroom. Not all content would be in the middle-school curriculum, but all content would be relevant either to data collected on the teachers' students via standardized assessment or to the teachers' action research projects.

As we planned the first iteration of the course, we were also cognizant of the challenges of the course structure itself. The course met for one week, eight hours per day. The teachers also typically stayed two to three hours after class to work on homework for the next day. Because of the limited contact time, we wanted to make in-class time as valuable as possible. We also realized that extended periods of lecture would not realistically keep teachers focused and on task. As a result, our class was designed to be activity-based, as opposed to lecture-based (Garfield 1993). The activities were designed to not only keep teachers engaged during class time, but to also help teachers discover major statistical concepts on their own. Some topics were still covered via lecture (e.g., sampling strategies, experimental design), but lecture time for these decreased with every subsequent course offering and was replaced by hands-on learning.

We used three main methods of assessment. First, daily homework was assigned. These typically consisted of 4-6 practice problems covering the day's content. The homework problems were assessed on a three-point scale: √+, √, √- for each of two levels: content and engagement (Table 2). Teachers were allowed to revise and re-submit homework if they desired. This grading scale was consistent with other M^2 courses, and so the teachers were accustomed to it.

Table 2. Daily Homework Grading Scale

Engagement Assessment Scale:	Content Assessment Scale:
√-: Modest (or no) homework submitted	√- : Solution needs to be improved
√: Satisfactory evidence of homework	√: Solution is essentially correct
√+: Work exceeds expectations	√+: Solution is correct and of high quality

Each daily homework set also included at least one "Habits of Mind" problem. These problems were designed to stretch their statistical thinking and reasoning skills beyond what was expected on the homework (http://scimath.unl.edu/MIM/habits.php). For example, one Habits of Mind

problem focused on experimental design and used an example involving standardized assessment of student achievement. The teachers were asked to read an excerpt from *Freakonomics* (Levitt and Dubner 2005) involving teachers in the Chicago Public Schools. The excerpt included answer strings from two teachers in sixth-grade classrooms in Chicago, and it was noted that one answer string was from a teacher who faked his/her student answers and the other was from a non-cheating teacher. Our teachers were asked to decide which teacher they thought was the cheater, and to provide reasoning. In addition, they were asked to design an experiment for the Chicago Public Schools to "catch" cheating teachers, under the constraint that they had the resources to retest only 120 classrooms.

At the end of the course, each teacher was asked to submit an End-of-Course Assignment. They had roughly 2 weeks after the completion of the course to submit their work, and the assignment had three components: (1) the Super Seven, (2) an end-of-course problem set and (3) a lesson plan. The Super Seven was to consist of a "final" solution to seven problems they worked on during the course. They were asked to select seven problems that would demonstrate what they accomplished/learned over the course, explain why they chose the problems, and provide reflection on their learning through those problems. The end-of-course problem set consisted of 15 new homework-type problems covering the gamut of course material. For the lesson plan, the teachers were asked to develop a statistical lesson that could be used in a class they teach, and the lesson needed to include the active participation of their students. The teachers were encouraged to think beyond lessons for math classes, and to consider potential lessons for social studies, science or other classes. Final grades in the course were based on the rubric, adapted from other M^2 Institute grading schemes, which can be found in the Appendix.

Assessment methods and grading rubrics were all in place before the first iteration of STAT 892 was offered during July 18-July 22, 2005. We used the text *Intro Stats*, 2nd edition (De Veaux, Velleman and Bock 2006). We had reviewed several textbooks, and this was chosen due to the short chapters by topic, the thorough step-by-step instructions, and the way this text approached the topics. In addition, we thought the text would be easy for the teachers to read on their own, which we believed would make it a helpful reference for the teachers at the conclusion of the course. Our initial plan was to cover the following topics:

- Day 1: Graphical displays of univariate data, measures of center and spread
- Day 2: Graphical and numerical summaries of bivariate data (i.e., simple linear regression and two-way tables), sampling strategies, experimental design
- Day 3: Probability, random variables, sampling distributions
- Day 4: Confidence intervals and hypothesis testing for one- and two-sample problems
- Day 5: Synthesis of topics via data collection/analysis activities

Admittedly, this was an ambitious schedule. As the week progressed, we realized that it was not realistic and adjusted the pace of the course as we were teaching it. In the next two sections we will describe how the week actually progressed, and what was modified in subsequent course offerings. An outline of the course and assignments utilized after several revisions and many iterations of the course may be found at
http://scimath.unl.edu/MIM/coursematerials.php#STAT892.

3. First Iteration

Course days all followed a similar structure. We began each day (except the first) with a homework discussion. Teachers were randomly divided into five groups prior to the start of the course, and each group was assigned a member of the instructional team as "Staff Advisor." The role of this grouping was primarily to review daily homework with their groups and answer any lingering questions from the previous topics. Because the groups and staff advisors met daily, this arrangement helped teachers get more individualized attention and gave them the opportunity to feel comfortable with at least one of the instructional team members. Each morning, the groups met to discuss homework questions before the assignments were turned in for grading, or any other concerns that their staff advisor could then bring forth to the entire instructional team. At that time, teachers were also asked to sign up to present a solution. Each teacher was expected to present at least one solution during the week. Homework solution presentations were sprinkled throughout the day to break up the topics and maintain student engagement.

During the remainder of the class day, one member of the instructional team would lead a class lecture, discussion, or activity while the remaining members of the team would either grade homework or circulate through the classroom to answer one-on-one questions, provide individualized instruction for struggling teachers, or more detailed explanation for teachers who had questions on a deeper level about that topic. We used a collaborative co-teaching model where the instructional team members all shared responsibility for being the lead instructor for specific topics throughout the course. Changing the face at the front of the room broke up the class day, as well as allowed the teachers to get to know and feel comfortable with all of the instructors. It also reinforced the idea that all instructors were equally equipped to answer questions or provide guidance, and gave the instructors breaks from being at the front of the room. In addition, the instructors had time to prepare and adjust instruction for later topics as needed.

New course topics were generally introduced through a data collection activity. For example, before starting graphical displays of data on Day 1, the teachers completed a handedness inventory (Gelman and Nolan 2002) to determine the extent to which each teacher was right- or left-handed. The resulting data was used to demonstrate the construction of graphical summaries like histograms, stem-and-leaf plots and boxplots. Class activities were also used to reinforce course concepts. For example, after the discussion of sampling strategies on Day 2, the teachers completed the Rolling Down the River activity (Doetsch, Flanagan-Hyde, Harrison, Tabor, and Tibeno 2000) to practice collecting simple and stratified random samples and to think about when one strategy might be more appropriate to use than another. After the discussion of experimental design, the teachers re-created Fisher's tea tasting experiment, to determine if they could tell the difference between regular and diet Dr. Pepper.

Each day ended with the teachers completing a daily evaluation. On the evaluation, they were asked to reflect on questions such as "What work was helpful today? Did at least one new insight grow out of today's sessions? If yes, what?" and "Did anything presented today affect your statistical thinking and/or your attitude toward statistics?" Additional comments, such as things

they liked and/or disliked or topics on which they needed more examples or instruction, were also encouraged. The instructional team reviewed the evaluations each day and made any necessary adjustments for the following day's plan.

Our initial pedagogical methods were focused on calculation, as was typical for a pre-GAISE classroom. While this mindset evolved over the various course offerings, we had varying levels of success with these methods in the initial course offering.

Most of the teachers had previous experience with topics toward the beginning of the class, such as summarizing both univariate and bivariate data. Many commented that they taught things like boxplots and simple linear regression in their own courses, and these went fairly smoothly. Day 3, which began with probability, was when it became apparent that our schedule was not realistic. The teachers became frustrated with the quantity of material being covered, and our pace slowed considerably. Sampling distributions, which we had predominantly motivated through simulation, were a major sticking point.

After observing the classroom dynamics and reading the Day 3 daily evaluations, we attempted to make adjustments before the final two days of the class (Table 3). Rather than glossing over sampling distributions and skipping to inference, we spent more time on them and used empirical sampling distributions to motivate the notions of interval estimation and the p-value as a measure of unusualness. Because we realized this would take up a lot of time (even though it was time well spent), we knew that some concepts would need to be dropped. We decided to skip many of the calculation details involved in one- and two-sample inference and to carry out all of the inference procedures using the TI-84. This allowed us to focus on interpretation and the logic behind the inference methods. In the end, we did not have Day 5 free to synthesize the course content, but adjustments made for subsequent iterations of the course allowed for this.

Table 3. Course Plan and Revisions during Cohort 1

Day	Initial Plan	Revised Plan	Modification
1	Graphical displays of univariate data, measures of center and spread	Summarizing univariate and bivariate data	None, teachers had experience with these topics
2	Graphical and numerical summaries of bivariate data (i.e., simple linear regression and two-way tables), sampling strategies, experimental design	Summarizing bivariate data, sampling strategies, experimental design	None, teachers had experience with these topics and were comfortable with definition material
3	Probability, random variables, sampling distributions	Probability and sampling distribution	Slower pace, teachers were frustrated with material
4	Confidence intervals and hypothesis testing for one- and two-sample problems	Sampling distribution, CIs, HTs, p-value	Focused on interpretations instead of calculations
5	Synthesis of topics via data collection/analysis activities	One and two-sample inference	No synthesis of course content for Cohort 1, adjusted for later groups

Two forms of evaluations were collected throughout the STAT 892 course. In addition to the daily evaluations previously described, the teachers completed an end of course evaluation that allowed them to constructively assess the course as a whole. On these evaluations, the teachers noted that our initial approach toward STAT 892 was fairly traditional, focusing more on calculations and formulas than on conceptual understanding. Some of the teachers in this first group of participants (Cohort 1) were frustrated with these calculations and expressed difficulties understanding and identifying relationships between various methods with the overwhelming amount of material presented. After the third day of the course, a teacher commented, "The mass quantity of information received didn't allow much time for reflection," while another wrote, "Slow down, please! Quality vs. Quantity. I'd rather understand a few things well as opposed to understanding a whole bunch of things poorly." The teachers also commented about the instructional team, and how they appreciated the team approach to sharing lecture duties and that someone was always available to help throughout the day or on homework in the evening.

Immediately following the completion of the course, the instructional team met to de-brief about our experiences and participant feedback from the course evaluations, and we determined that much needed to be changed for the next iteration of the course based on our initial experiences. This course re-structuring and re-envisioning occurred after each subsequent iteration of the course, and the changes we made with each cohort of teachers are described in the next section.

4. Subsequent Iterations

Immediately following the conclusion of Cohort 1's course, the instructional team met to discuss changes for the next cohort, beginning that fall. After the experience with Cohort 1, we decided to transition the course to focus more on statistical thinking and less on mechanics and calculations. A major decision was made regarding course content: remove discussion of the binomial probability formula. We decided that focusing on the conceptual nature of a binomial experiment (e.g., flipping a coin) was sufficient for the teachers to understand probabilities of binary outcomes rather than emphasizing the binomial probability formula itself. While this change was difficult for us to make as instructors of a traditional introductory course, it was the beginning of other changes leading to emphasizing concepts over mechanics. As recommended by the GAISE Report (Aliaga et al. 2005), course lectures, activities, and assignments were restructured to put the emphasis on big picture statistical concepts and not on calculations by hand. One or two examples of calculations by hand were shown for each topic so teachers could understand how the statistical procedure worked, and then the same examples were done using the TI-84 calculators. This approach had been used for some topics in Cohort 1, but as a result of running out of time rather than a shift in instructional approach. Technology was utilized much more in the later cohorts, which freed up time for statistical reasoning and thinking. Instead of furiously writing down each step of each problem, the teachers could spend more time thinking through the problems and concepts.

All iterations of STAT 892 covered the same material one would find in an undergraduate non-calculus based introductory statistics course over periods of time varying from one week to a semester. There were many changes made between cohorts, both in the structure of the course and in the delivery method (Table 4). Cohort 1's course was structured in a traditional format focusing on the mechanics of summary statistics and statistical inference. Cohort 1 took the course over one week in the summer. Class was from, 8:00 a.m. to 5:00 p.m., with several hours of homework each night. Although there were several hands-on activities during the day, many of the overarching statistical concepts were lost in the focus on mechanics and calculations. Cohort 2's course was offered partially in-person and partially online. The teachers came together for two days in the fall where they had the same lectures, activities, and assignments as the first two days of Cohort 1's course. Then, they completed the remainder of the course online from their home locations over the rest of the fall semester. The online portion utilized the Blackboard© course management system, and the teachers were assigned to discussion groups to work together on homework and discuss content. Cohorts 3-5's classes were all taught face-to-face during the summer. Cohort 3's course met for two weeks from 8:00 am to noon, and Cohorts 4 and 5's classes had the same one week structure as Cohort 1.

Table 4. Changes in Format and Focus between Cohorts

Cohort	Format and Delivery Mode	Major Changes from Previous Offering
1	5 days 8:00 a.m. to 5:00 p.m.	Initial offering: traditional format, focus on mechanics and calculation
2	2 days 8:00 a.m. to 5:00 p.m. Remainder of course online over a semester	Started shift to conceptual focus; Teachers would collect data on their own, then instructor would compile; Utilized discussion board
3	10 days 8:00 a.m. to 12:00 p.m.	Continued shift to conceptual focus; incorporated data analysis project
4-5	5 days 8:00 a.m. to 5:00 p.m.	Continued emphasis on conceptual focus; continued and expanded data analysis project; incorporated more real data related to students and education; more time on good and bad graphs

Many of the hands-on activities used in Cohort 1's course involved generating data (rolling dice, flipping a coin, etc.) to build and visualize distributions. While these activities were still used with Cohort 2, it was more difficult due to the online format. For example, instead of rolling a 10-sided die, teachers had to simulate this data using their TI-84 calculators and report their results. A course instructor would then compile the data and build the distributions for discussion. Much of the effectiveness of building a distribution using data generated in a class setting was lost. Although homework problems, course material, and statistical ideas were discussed in groups using the online discussion board, the teachers did not seem to grasp concepts as well in their isolated settings. In Cohort 2, teachers did not appreciate the online format of the course. Through this form of instruction, several teachers struggled with grasping the statistical concepts. One teacher remarked, "I had difficulty understanding the theory behind much of the course. I can follow the examples and 'get the answer' but understanding 'WHY' it works is more difficult. I think a 'face to face' class would have helped." Overall, classroom interaction was an important component missing (but desired) in the course to help solidify teachers' understanding of the concepts. One teacher commented, "It just did not connect for me. I need interaction to process...I have <u>much</u> stuff rattling around, but it is not jelling. I don't know when to use what, or why." While this approach may be more successful now with recent developments in technology and literature in online teaching and learning, at the time and based on evaluations from this Cohort, the program director decided to go back to the face-to-face setting for future offerings of this course.

The other major course change occurred between Cohorts 2 and 3: we incorporated a data analysis project. We reserved four hours the second to last day of the class for teachers to work in groups to analyze a dataset. About one hour was spent giving a description of the dataset and variables and demonstrating how to use Excel to do the same things they had learned to do using the calculators. Teachers then worked with their group members to come up with a set of

questions they wanted to answer using the data. They were responsible for utilizing descriptive statistics, appropriate graphs, tests or confidence intervals to answer their research questions and for preparing a presentation of their results. They were able to finish the presentation for that evening's homework and presented their questions and results to the class on the last day. This allowed teachers an opportunity to synthesize the information and demonstrate understanding, an opportunity that was missed due to time in Cohort 1. Through the daily and end of course evaluations, we were able to gauge teachers' perceptions about these course changes, and we used the teachers' feedback to inform each new iteration of STAT 892.

Considerably more emphasis was placed on conceptual understanding of statistical concepts and procedures in the course for Cohort 3 than had been previously placed in STAT 892 for Cohorts 1 and 2. Cohort 3 teachers valued the inclusion of the data analysis project, particularly because they were able to conduct unguided explorations with a real world data set. A teacher mentioned, "It was great to have us analyze our data with no directions just use what we know." Cohort 3 teachers thought this was a meaningful experience and suggested using it in future courses. A teacher recommended, "Continue with the projects and real data. I especially like the data analysis and presentations at the end as some groups tested something I never thought of testing." Overall, the project was a valuable experience; teachers enjoyed the activity and were able to discover the challenges associated with analyzing real data. After exploring the data, a teacher observed "that data doesn't always give you what you expect/want." Teachers even recognized the connection between the course content and the requirements for the action research project: "[STAT 892] made me understand the statistics that we will need to do our [action] research." This connection was vocalized more in the Cohort 3 course evaluations than any other previous cohort, and teachers in Cohorts 4 and 5 had similar observations and comments.

5. Discussion & Conclusion

Even through multiple iterations across five cohorts, this course was one of the favorites among the teachers. One teacher wrote, "This was my favorite MIM [Math in the Middle] course so far. It was applicable to [the] action research project and I loved that another aim was to make us more informed consumers of statistical data. Activities and assignments are also applicable in our classrooms." The course's success, even in the early stages, can partially be attributed to the care in planning and the willingness to adapt by the instructional team. After each cohort, we carefully considered the evaluation comments and the impressions from the week and made revisions accordingly. Although there is still room for improvement, by the end of Cohort 4, the team finally felt we had successfully combined recommendations in statistics education with the needs of the audience we were serving to conduct a class with the right content, level, order of topics, activities, etc., and we were able to retest the same structure with Cohort 5. By the end of the course, teachers not only appreciated the applicability of the course to their own research projects, but also to their own classrooms. Still, when planning an intensive course for a specialized audience, we have six specific recommendations.

First, carefully consider the audience for the course and their purpose in taking the course. Then, carefully consider what your goals are for the course, what you want the students to come away with, and make every effort to come up with creative ways to combine the two. In our course,

teachers were there to better understand mathematical and statistical concepts, but what we learned is that they have a strong desire to be able to understand and use the statistical information they are presented (e.g., proficiency standards, testing results) and they are always looking for ideas to use in their own classes. Several teachers commented that ours was one of the first courses where they did lesson plans and activities that could be directly applied to their own teaching. Teachers are very busy, but want to incorporate new ideas. Making activities they can use directly in their classes and the end of course lesson plan left teachers feeling like they came away with something useful and concrete. They received both content knowledge and engaging statistical learning activities they could easily use with their own students without added effort during the school year. We also used lesson plans from previous cohorts (with permission) to demonstrate concepts in future cohorts. Teachers appreciated this, and it let them know we valued their work, and that their end of course lesson plans could be utilized by themselves and serve as resources for others.

Second, with a specific group, tailor the content, examples, and activities to that audience as much as possible instead of reusing old, generic examples. Make efforts to tie all material to the purpose of the course, and communicate those ties that are not obvious to the audience. For example, when talking about z-scores and percentiles, we made a point to incorporate how it relates to interpreting standardized testing results.

Third, for courses delivered in a short time period, cut out extraneous material that doesn't relate directly to the goals or purpose of either the program, specific course, or the audience. After the first cohort, we made the decision to remove the binomial formula. For a group who had taught pre-GAISE introductory statistics, this was a difficult decision. In the end, it was a good one. Teachers did not need the formula to understand the concepts of binomial probability, and it saved a lot of time that we thought better spent on synthesizing information.

Fourth, as much as possible, be prepared for the personality of the group. This is another example of where the Master Teacher was an invaluable part of the instructional team. She helped to prepare us ahead of time that this would be a highly motivated, detail oriented group, with a sense of needing perfection in their answers. Even though we had this information, we were still not fully prepared until we taught this group. They were very concerned with details, such as how many digits to round their answer (and did not like our "2 or 3" response), and at first were very uncomfortable with not always having a black-and-white, right or wrong, answer. The teachers in the class often felt much more comfortable in expressing their concerns or frustrations to the Master Teacher as she was seen as more of a peer rather than an instructor. In Cohorts 3-5, the Master Teacher was a graduate of the program so she had also experienced the stresses that are inherent to an intensive program such as this, which helped us be more aware and more sensitive to this issue.

Fifth, if possible, use an experienced instructional team. We attribute much of the success of this course to the collaborative co-teaching approach that was used and think this is critical when working with a large group during an intensive, concentrated period of time. Each member of the team brings his or her own strengths, and it also allows for providing different viewpoints and insights. It also allows for division of labor to help accomplish goals and keeping energy high during an intensive instructional week. In addition, we learned that sometimes all it takes is a

different face in front of the room to break up the day, or an alternate way of explaining a concept to make the light bulb turn on.

Finally, always be open to change and adaptation. Changes we made both during each cohort based on specific group needs, and between the cohorts that had a big impact include changing from focusing on mechanics to focusing on statistical thinking and reasoning, having our Master Teacher member of the instructional team be a program graduate, and incorporating a data analysis project toward the end of the course. Things we continually changed, and would still work on for subsequent iterations would be the schedule. Day 3 of the week continued to be the most difficult day for the teachers and at first they had difficulty making the connections. This was also the day where topics got away from anything they had previously learned. We incorporated reviews and a summary table in Cohorts 4 and 5 which helped tie concepts together. The other main item to work on is the balance between activities and lecture. We did modify this throughout the cohorts and got better at the balance, but it still felt like this could be improved.

Overall, we felt the class was very successful in meeting the goals we had set. Based on evaluations the teachers seemed to agree, and the course improved with each delivery. It was well-received by the teachers, and they appreciated the care that went into tailoring the class to not only their needs as students, but also their needs as professionals and educators.

6. The Future of STAT 892

After STAT 892 was taught five times to Math in the Middle cohorts, it was incorporated on a semi-permanent basis into the Nebraska Math and Science Summer Institutes (NMSSI), week-long graduate courses for in-service teacher professional development. The course was adapted to fit a high school teacher audience, and has been taught two times each for middle level and high school teachers as part of the NMSSI (2009-2012). The middle level version of the course has been archived on the Math in the Middle website (http://scimath.unl.edu/MIM/coursematerials.php#STAT892) so other instructors can access our materials. Additionally, in 2012, the course was paired with a statistics pedagogy course designed by two master teachers who teach high school statistics courses. These two master teachers attended STAT 892 in the morning in order to cohesively integrate the STAT 892 content into their pedagogy lessons in the afternoons.

In 2013, the university received funding for two additional cohorts of Math in the Middle teachers, with all participants from a single, high-need school district. In summer 2014 this course was taught to Math in the Middle teachers, and will be again in summer 2015, giving us yet another opportunity to further the evolution of the course and enhance the statistics knowledge of educators in Nebraska.

Appendix

Final Grading Rubric from course syllabus

Part of our responsibility as the instructional staff is the assessment of participants' achievement in each Math in the Middle course. We recognize that teacher-participants are drawn from different grade levels, have different certifications to teach mathematics, and have different educational backgrounds with respect to previous opportunities to learn statistics. Thus, we believe it is appropriate to have an assessment system that values effort, teamwork, progress in learning statistics, and the development of teaching skill and leadership development.

Grade	*Expectations and typical characteristics of achievement at that level*
A+	*The grade of A+ is honorific and will be fairly rare. It is evidence that the instructors have special admiration for the participant's achievements in the course.*
A	*Achievement beyond the level needed to earn the grade of A-. Especially important will be evidence that the teacher has a good command of the statistics studied in the course, the ability to transfer statistics learned into the teacher's classroom, and progress in developing the ability to be a leader of one's peers.*
A-	*Achievement beyond the level needed to earn a grade of B+. In particular, there should be clear evidence of significant progress in learning statistics and in learning about issues that impact teachers' ability to help their students learn statistics. Activities that can contribute to earning the grade of A- will include going beyond the minimum expectations for homework during the summer institute.*
B+	*Regular class attendance, active participation, assignments submitted on-time, supportive and helpful to peers, admirable effort to complete assignments, evidence of good progress in learning statistics and transferring new statistics learned to the teacher's classroom.*
B	*Regular class attendance, reasonable participation, most assignments submitted on-time, cooperative with peers, reasonable effort to complete assignments, to learn statistics and to strengthen teaching skills.*
B-	*A grade below B is a statement that the instructors do not believe that the teacher made a reasonable effort to use the opportunity provided by the Math in the Middle Institute to develop into a stronger teacher. Evidence may include one or more of the following traits: attendance problems, uncooperative behavior, failure to submit assignment(s), habitual tendency to submit assignments late, or performance on assignments that indicates an inadequate effort to learn statistics and to strengthen teaching skills.*

Acknowledgments

References

Abrahamson, D., Janusz, R. H., and Wilensky, U. (2006), "There once was a 9-block…A middle school design for probability and statistics," *Journal of Statistics Education* [online], 14, 1. Available at http://www.amstat.org/publications/jse/v14n1/abrahamson.html

Alder, J., and Davis, Z. (2006), "Opening another black box: researching mathematics for teaching in mathematics teacher education," *Journal for Research in Mathematics Education*, 37, 4, 270-296.

Aliaga, M., Cobb, G., Cuff, C., Garfield, J., Gould, R., Lock, R., Moore, T., Rossman, A., Stephenson, B., Utts, J., Velleman, P., and Witmer, J., (2005), *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: College Report,* Alexandria, VA: American Statistical Association.

Batenero, C., Godino, J.D., and Roa, R. (2004), "Training teachers to teach probability," *Journal of Statistics Education* [online], 12, 1. Available at http://www.amstat.org/publications/jse/v12n1/batanero.html

Carmichael, C., Callingham, R., Watson, J, and Hay I. (2009), "Factors influencing the development of middle school students' interest in statistical literacy," *Statistics Education Research Journal*, 8, 1, 62-81.

Cobb, G., W. (1992), Teaching statistics. In Lynn A. Steen (Ed.), Heeding the call for change: Suggestions for curricular action (MAA Notes No. 22, pp. 3-43).

Cobb, G. W. (1993), "Reconsidering statistics education: A National Science Foundation Conference," *Journal of Statistics Education* [online], 1, 1. Available at http://www.amstat.org/publications/jse/v1n1/cobb.html

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.

Conference Board of the Mathematical Sciences (2001), *The Mathematical Education of Teachers,* Providence, R.I. 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, R.I. and Washington, DC: American Mathematical Society and Mathematical Association of America.

De Veaux, R. D., Velleman, P. F., and Bock, D. E. (2006), *Intro Stats*, 2nd Edition, Boston, MA: Pearson Education, Inc.

Doetsch, C., Flanagan-Hyde, P., Harrison, M., Tabor, J., and Tibeno, C. (2000), "An exercise in sampling: Rolling down the river," *Student Activities on Experimental Design and Web-based Resources for Teaching* [online]. Available at http://courses.ncssm.edu/math/Stat_inst01/PDFS/river.pdf

Franklin, C., Kader, G., Mewborn, D. S., Moreno, J., Peck, R., Perry, M., and Scheaffer, R. (2007), *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework,* Alexandria, VA: American Statistical Association.

Garfield, J. (1993), "Teaching Statistics Using Small-Group Cooperative Learning," *Journal of Statistics Education* [online], 1, 1, Available at http://www.amstat.org/publications/jse/v1n1/garfield.html.

Garfield, J. (2000), "An Evaluation of the Impact of Statistics Reform," Final Report for National Science Foundation project REC-9732404.

Garfield, J. (2002), "The challenge of developing statistical reasoning," *Journal of Statistics Education* [online], 10, 3. Available at http://www.amstat.org/publications/jse/v10n3/garfield.html

Garfield, J., and Everson, M. (2009), "Preparing teachers of statistics: A graduate course for future teachers," *Journal of Statistics Education* [online], 17, 2. Available at http://www.amstat.org/publications/jse/v17n2/garfieldpdf.html

Gelman, A., and Nolan, D. (2002), *Teaching statistics: a bag of tricks*. Oxford: Oxford University Press.

Heaton, R. M., Lewis, W.J., and 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., and Smith, W. M. (2013), "The Math in the Middle Institute: Strengthening middle level teachers' mathematical and pedagogical capacities," in *Resources for Preparing Middle School Mathematics Teachers,* eds. C. Beaver, L. Burton, M. Fung, and K. Kruczek, Mathematical Association of America Note Series, pp. 47-58.

Heaton, R. M., Lewis, W. J., Homp, M. R., Dunbar, S. R., and Smith, W. M. (2013), "Challenging and rigorous yet accessible and relevant mathematics courses for middle level

teachers," in *Resources for Preparing Middle School Mathematics Teachers,* eds. C. Beaver, L. Burton, M. Fung, and K. Kruczek, Mathematical Association of America Note Series, pp. 181-201.

Leavy, A. M. (2010), "The challenge of preparing preservice teachers to teach informal inferential reasoning," *Statistics Education Research Journal*, 9, 1, 46-67.

Lee, H. S., Angotti, R. L., and Tarr, J. E. (2010), "Making comparisons between observed data and expected outcomes: Students' informal hypothesis testing with probability and simulation tools," *Statistics Educaiton Research Journal*, 9, 1, 68-96.

Lesser, L. M., and Kephart, K. (2011), "Setting the tone: A discursive case study of problem-based inquiry learning to start a graduate statistics course for in-service teachers," *Journal of Statistics Education* [online], 19, 3. Available at www.amstat.org/publications/jse/v19n3/lesser.pdf

Levitt, S. D., and Dubner, S. J. (2005), *Freakonomics: A rogue economist explores the hidden side of everything,* William Morrow Paperbacks.

Metz, M. L. (2010), "Using GAISE and NCTM Standards as Frameworks for Teaching Probability and Statistics to Pre-service Elementary and Middle School Mathematics Teachers," *Journal of Statistics Education* [online], 18, 3. Available at http://www.amstat.org/publications/jse/v18n3/metz.pdf.

Moore, D.S. (1997), New pedagogy and new content: The case of statistics.*International Statistical Review, 65*, 123-137. http://www.stat.purdue.edu/~dsmoore/articles/

National Council of Teachers of Mathematics (1989), *Curriculum and Evaluation Standards for School Mathematics*, Reston, VA: NCTM. (http://www.nctm.org/standards/content.aspx?id=26629)

National Council of Teachers of Mathematics (2000), *Principles and Standards for School Mathematics,* Reston, VA: Author. http://www.nctm.org/standards/content.aspx?id=318.

Rossman, A., Chance, B., and Medina, E. (2006), "Some Important Comparisons between Mathematics and Statistics, and Why Teachers Should Care" in *Thinking and Reasoning With Data and Chance: Sixty-eighth Yearbook,* eds. G. F. Burrill and P. C. Elliott, Reston, VA: National Council of Teachers of Mathematics, pp. 323-333.

Scheaffer, R. L. (1986), "The Quantitative Literacy Project," *Teaching Statistics, 8(2).*

Scheaffer, R. L. (1991), "The ASA-NCTM Quantitative Literacy Project: An Overview,"*Proceedings of the 3rd International Conference on Teaching Statistics*, pp. 45-49.

Scheaffer, R. L., and Jacobbe, T. (2014), "Statistics education in the K-12 schools of the United States: A brief history," Journal of Statistics Education [online], 22, 2, Available at http://www.amstat.org/publications/jse/v22n2/scheaffer.pdf.

Schifter, D. (1993), "Mathematics process as mathematics content: A course for teachers," *The Journal of Mathematical Behavior*, 12, 3, 271-283.

Schifter, D., and Fosnot, C. (1993), *Reconstructing Mathematics Education: Stories of Teachers Meeting the Challenge of Reform*, New York, NY: Teachers College Press.

Simon, M. A., and Schifter, D. (1991), "Towards a constructivist perspective: An intervention study of mathematics teacher development," *Educational Studies in Mathematics*, 22, 4, 309-331.

Simon, M. A., and Schifter, D. (1993), "Toward a constructivist perspective: The impact of a mathematics teacher inservice program on students," *Educational Studies in Mathematics*, 25, 4, 331-340.

Simon, M. A. (1994), "Learning mathematics and learning to teach: learning cycles in mathematics teacher education," *Educational Studies in Mathematics*, 26, 1, 71-94.

Smith, T. M., and Hialmarson, M. A. (2013), "Eliciting and developing teachers' conceptions of random processes in a probability and statistics course," *Mathematical Thinking and Learning: An International Journal*, 15, 1, 58-82.

Watson, J. M. (2008), "Exploring beginning inference with novice grade 7 students," *Statistics Education Research Journal*, 7, 2, 59-82.

Kendra K. Schmid
University of Nebraska Medical Center
984375 Nebraska Medical Center
Omaha, NE 68198-4375
(402) 559-8117
kkschmid@unmc.edu

Erin E. Blankenship
University of Nebraska-Lincoln
Department of Statistics
340 Hardin Hall, North Wing
Lincoln, NE 68583-0963
(402) 472-7398
erin.blankenship@unl.edu

April T. Kerby
Winona State University
Department of Mathematics & Statistics

P.O. Box 5838
Winona, MN 55987
(507) 457 – 2354
akerby@winona.edu

Jennifer L. Green
Montana State University
P.O. Box 172400
Bozeman, MT 59717-2400
(406) 994-3601
jgreen@montana.edu

Wendy M. Smith
University of Nebraska-Lincoln
Center for Science, Mathematics, and Computer Education
251 Avery Hall
Lincoln, NE 68588-0131
(402) 472-7259
wsmith5@unl.edu