2011

Perspectives on Deepening Teachers' Mathematics Content Knowledge: The Case of the Oregon Mathematics Leadership Institute

Libby Knott
Washington State University

Martha VanCleave
Linfield College

Perspectives on Deepening Teachers' Mathematics Content Knowledge: The Case of the Oregon Mathematics Leadership Institute

July 2011

Libby Knott
Martha VanCleave

Work of the Oregon Mathematics Leadership Institute MSP supported by the National Science Foundation (Grant number 0412553).

Writing of this case was supported by the Math and Science Partnership Knowledge Management and Dissemination Project, funded by the National Science Foundation (Grant number 0445398) under the direction of Iris R. Weiss of Horizon Research, Inc., and Barbara A Miller of Education Development Center, Inc.

These writings do not necessarily reflect the views of the National Science Foundation.

Abstract
The Oregon Mathematics Leadership Institute (OMLI) project served 180 Oregon teachers, and 90 administrators, across the K-12 grades from ten partner districts. OMLI offered a residential, three-week summer institute. Over the course of three consecutive summers, teachers were immersed in a total of six mathematics content classes– Algebra, Data & Chance, Discrete Mathematics, Geometry, Measurement & Change, and Number & Operations—along with an annual collegial leadership course. Each content class was designed and taught by a team of expert faculty from universities, community colleges, and K-12 districts. Each team chose a few "big ideas" on which to focus the course. For example, the Algebra team focused on algebraic structure and properties of the concept of a group, while the Data & Chance team centered their activities on the exploration of ideas of central tendency and variation using statistics and data analysis software packages. The content in all of the courses was addressed through deep investigation of the mathematics of tasks that had been selected and adapted from resources for K-12 mathematics classrooms. In addition to mathematics content, the courses were designed with specific attention to socio-mathematical norms, issues of status differences among learners, and the selection and implementation of group-worthy tasks for group work. The faculty attended sessions grounded in the work of Elizabeth Cohen on strategies for working with heterogeneous groups of learners (Cohen, 1994; Cohen et al, 1999) which was central to the OMLI design and implementation. Institute faculty modeled these strategies in the Institute classrooms and made their moves as transparent as possible, so that the teachers would be able to grapple with these strategies during the Institute and plan for implementation in their own classrooms. The Data & Chance course also modeled uses of technology in instruction using *Tinkerplots*. Generalization and justification were emphasized as mathematical ways of learning and knowing, and institute faculty conducted classroom discussions that intentionally modeled pushing for generalization and justification.

Overview of the project

The Oregon Mathematics Leadership Institute (OMLI) project included teachers and administrators from ten partner districts, serving teachers across all grades from kindergarten to 12th grade. Each participating district was invited to assemble school or district teams. School teams typically consisted of two teachers and a principal; district teams consisted of two teachers and a district level administrator. The NSF funding was intended for 100 teachers and 50 administrators; additional funding provided by the state of Oregon's Mathematics and Science Partnership program and district support made it possible for OMLI to serve approximately 180 teachers and 90 administrators.

The professional development program was offered through a residential, three-week summer institute, designed as a comprehensive program over the course of three consecutive summers. These sessions were held jointly on the Oregon State University campus and at a local middle school. The mathematics content knowledge classes for teachers were offered along with collegial leadership classes for both teachers and administrators. The collegial leadership courses were designed to provide the participants with both the skills and confidence needed to build professional learning communities with their colleagues in their own schools and districts.

Six mathematics content classes were offered each summer – Algebra, Data & Chance, Discrete Mathematics, Geometry, Measurement & Change, and Number & Operations. Each course was designed and taught by a team of four expert faculty from institutions of higher education, Community Colleges, and K-12 districts, in consultation with K-12 teachers, district curriculum specialists and administrators. Each of the six mathematics content courses was offered in two sections, with each section offered twice daily. The classes were small, serving about 15 - 20 participating teachers per class. Each summer, participants took two mathematics content classes and a collegial leadership class. Whenever possible, participants from the same district took the same classes each summer. The leadership class was held in the afternoons, in two back-to-back sessions. After their two morning content classes, the participants attended one of the leadership class sessions, with the remainder of the afternoon devoted to homework and study time. Institute faculty were available on a rotating basis during these sessions, and homework was coordinated so that appropriate faculty were available on the days when their assignments were due.

Course Design and Instructor Preparation

Prior to the start of the Institutes, and to aid with the choice of topics and the design of the courses, the Institute faculty teams had met with representatives from the participating schools to discuss the particular needs of their schools. This information, together with the content objectives addressed in the state standards, was used as the starting point for the course design, which was completed during the six months prior to the first summer. In subsequent academic years, the courses were revised and streamlined to incorporate lessons learned during the previous summer's institute.

Faculty teams were given the freedom to include the material they believed would best address these needs. These teams had access to information and workshops about "best practices" in instructional methods, mathematical knowledge for teaching and pedagogical content knowledge, the design of group-worthy tasks, group protocols and much more. Each team chose a few "big ideas" on which to focus the course. The Algebra team focused on the central idea of the algebraic structure and properties of the concept of a group, exploring this idea through the symmetries of geometric shapes. The Data & Chance team centered their activities on the use of *Tinkerplots*™ and *Fathom*™ (Key Curriculum Press, 2005a, 2005b) using these statistics and data analysis software packages to explore ideas of central tendency and variation. The Discrete Mathematics team focused their activities on graph theory, combinatorics, cardinality, patterns and cryptography. The Geometry team focused their activities on comparing and contrasting properties of Euclidean geometry with those of non-Euclidean (spherical) geometry. The Measurement & Change team built their curriculum around the concept of change in one, two, and three dimensions and how to measure it, exploring length, area and volume. The Number and Operations team designed their activities to provide exploration in prime numbers and factoring, fractions and ratios, and the arithmetic properties of integers.

Each participant enrolled in two of the six courses each of the three summers. The faculty from the paired courses worked in consultation with each other, planning complementary activities and homework assignments, and in some cases planning joint final projects for the participants. Two design principles were consistently emphasized: (1) mathematical discourse specific to the concepts of justification and generalization was modeled in each of the content classes, so as to make these concepts and the discourse surrounding them explicit; and (2) the faculty teaching the courses modeled an activity-based instructional style using small groups and group protocols, and emphasizing social and socio-mathematical norms.

The Institute courses addressed the content knowledge needs of the participating teachers, but offered far more than a typical content course. The K-12 mix of teachers was a challenge for the Institute faculty at first. They had to learn to deal with issues of status that arose immediately due to the interactions of elementary teachers and high school teachers. The faculty had received training in the establishment of social and socio-mathematical norms in the classroom, issues of status, and the design of group-worthy tasks and the use of group work. All of these issues were addressed in the classes, with Institute faculty both modeling these strategies and making their moves as transparent as possible, so that the teachers would be able to grapple with these strategies during the Institute and plan for implementation in their own classrooms.

Professional Development: The Data & Chance Curriculum

All the OMLI faculty teams took as their starting point for curriculum materials and rich problems for all the OMLI courses, standards-based middle and high school curriculum materials. In the Data and Chance course middle and high school activities were used to provide access to all participants while providing opportunities to deepen the content knowledge and challenge participants with prior conceptual understanding. While some activities may have been suitable for classrooms of the participants, the emphasis was on developing deep conceptual understanding and not on providing activities for participants to use in their own classrooms. The teams identified the “big ideas” in their curriculum strand, and adapted materials as necessary to provide activities with multiple access points for K-12 teacher participants. The OMLI Data & Chance course focused on exploring data, investigating chance, and connecting data and chance, with daily 110-minute class sessions over the three-week summer institute. Planning for the course began with a symposium for participants and faculty during which the participants clearly communicated their desire for content instruction delivered utilizing the classroom practices they were working to implement in their own classrooms. Particularly desirable practices included opportunities for meaningful student mathematical discourse, establishment of socio-mathematical norms, attention to status issues, and group work with truly group-worthy activities. The goal was to design tasks in such a way as to not give any particular advantage to the person who might know an algebraic formula, a particular algorithm, or a specialized piece of mathematical information. The course designs would also make sure that all teachers had time to read, reflect on and work on each activity, and all attempts and viewpoints were listened to and valued.

The faculty team for the Data & Chance course included a middle school master teacher and three college faculty: a mathematics educator from a large state university, a mathematician from a small liberal arts university, and a mathematics educator from a small liberal arts college. The team members did not have experience team teaching and in some cases were not familiar with the desired instructional practices; so working together to design and present the course required that they participate in several OMLI faculty planning retreats that included team-building, review and discussion of literature on instructional practices, and time to work on course development. Attention to status issues proved to be especially important in the OMLI project because all courses were designed and implemented for K-12 teachers learning together. Because group activities were central to the course, a three-session course on designing group work was provided.

The development and revision of the Data & Chance course was informed by the research on teaching statistics in the Guidelines for Assessment and Instruction in Statistics Education (GAISE) report (Franklin et al, 2005). The National Council of Teachers of Mathematics *Navigating through Data Analysis* and *Navigating through Probability* series (Bright, Brewer, McClain, & Mooney, 2003; Bright, Frierson, Tarr, & Thomas, 2003; Burrill, Franklin, Godbold, & Young, 2004; Chapin, Kozoil, MacPherson, & Rezba, 2002; Shaughnessy, Barrett, Billstein, Kranendonk, & Peck, 2004; Sheffield et al, 2002) were used as texts for the course and some activities were adapted from these materials. Each lesson began with a focus question from which objectives and core ideas were developed and activities were selected or created.

The first week of the Data & Chance course focused on establishing socio-mathematical norms and developing a "data detective" philosophy prompted by looking at a data set through the lens of the questions "What do you notice?" and "What do you wonder about?" Activities provided opportunities to explore a variety of visual representations and numerical descriptions of data, which could then be used in comparing data sets. The ideas of variation and reliability in measurements were investigated and discussed. Relationships between numerical and categorical data were explored graphically with discussion of the idea of correlation and caution about the assumption of causation.

During week two the focus was on probability, beginning with the language of chance and defining probability as the long-range relative frequency of an event. Activities provided opportunities to explore experimental as well as theoretical probabilities of events, multistage experiments, conditional probabilities, independent and dependent events, probability distributions, sample size, and expected value. Real world problems were explored using simulations that were analyzed and discussed.

Week three pulled together the work of weeks one and two with activities that used data to make predictions and base decisions on data and simulation. The idea of confidence in decision-making was highlighted through discussion of confidence intervals and the idea of testing a hypothesis.

The course culminated with presentation of statistical posters prepared following the guidelines of the Americans Statistical Association Poster Competition (American Statistical Association, 2010), through which participants demonstrated their understanding of data analysis and decision-making.

Emphasis on variation was central to the design of the Data & Chance curriculum. The course began with an exploration of the time between eruptions of the *Old Faithful* geyser. This data set exemplifies the need to consider more than the measures of center: mean, median, and mode. When the data are presented visually, especially in a plot over time, the pattern of variation is revealed and the real story of the data can be understood. This focus on variation is wellillustrated in the *Body Measurements* lesson, and the *Chip Sampling from Known and Unknown Mixtures* lessons. A description of each lesson and its implementation follows.

The Body Measurements Lesson

In the *Body Measurements* lesson, participants investigated sources of variation in data by physically gathering data on body measurements of each of four instructional staff. Participants made observations about the data collected by the class, paying particular attention to the variation in the data and possible sources of the observed variation. Participants, as a group, discussed and implemented strategies in the data gathering process to control for variation and the process was repeated.

The focus question for this activity was, "What is meant by the variability of data?" The activity was designed with the objective for participants to recognize different sources of variation including those that can be minimized or controlled (noise within measurements) and those that are an integral part of the data (signature of variation). The lesson addressed the core ideas: sources of variation, data production/collection, and reliability controls.

The scenario for this activity was, "*When people buy clothes, they often get things that fit approximately, but there are cases where clothes must fit exactly, such as astronauts' suits. We are going to take some measurements to see what we would report to the astronaut tailors for different parts of an astronaut suit.*" The participants were then divided into four heterogeneous groups with teachers from different grade levels in each group. Having teachers from different grade levels in each group provided opportunities for them to learn from one another. High school teachers might clarify some misconceptions held by teachers from earlier grades while learning what would be expected of students at different grade levels. Knowing what teachers at each level would expect from their students if the activity were to be implemented in their classrooms helped clarify the continuum of data and chance instruction across all grade levels.

The four instructors were stationed around the room. Participants were instructed to take a single measurement at each of the measuring stations: *armspan*, *headspan*, *forearm length* and *footlength*. It was critical that participants be given no additional instruction initially because the ambiguity in measurement criteria at this point was part of the learning trajectory for this lesson. They were not told how they should take the measurements, nor was there any clarification of what was meant by any of the quantities. Participants were given cloth measuring tapes approximately 60 inches in length, deliberately not long enough to measure the full length of an armspan. After each measurement, participants recorded their measurements on a dot plot at the measuring station. Each group rotated through the stations until they had completed a measurement at each station. Not surprisingly, since no definitions were given for the measurements, some participants deliberately took measurements in unorthodox manners, such as headspan from the crown of the head to the point of the chin. Anticipating possible push-back from participants who failed to perceive the value of the activity highlights the importance of being aware of and prepared to confront issues of status and attempted power plays. In this lesson, clearly articulating the objective, to explore variation in the data, is essential to counteracting this situation. Participants needed to recognize that their students would not have experience with the idea of variation and would require activities such as this one to gain an understanding of the importance of the concept.

After all the measurements were completed, the resulting dot-plots were displayed. Private think time was provided for participants to examine the graphs and record their observations concerning the data for these repeated measures. After a few minutes, one of the instructors led a class discussion focusing on what the dot-plots revealed. Typical discussion prompts included:

- What do the graphs reveal about body measurements?
- What do you notice? What do you wonder about?
- Do the graphs look the way you would expect them to?
- Can you explain why the graphs display the kinds of variation seen?
- Do all the measurements display the same amounts of variation? If not, what might account for the differences in variation?

Participants were then asked what could be done to help ensure that repeated measures would be closer together. Each of the groups was then assigned one attribute for which they created at most three measurement criteria. These criteria were things such as "always measure the right foot," or "measure the foot with the shoe and sock removed." Each group recorded their criteria on a poster, which they then shared with the whole class. Individuals were allowed to ask clarifying questions in order to understand the criteria but not to suggest alterations. Participants repeated the measurement exercise, following the criteria established by the particular group. Again, dot-plots were created to display the data. Participants compared the posted dot-plots for the two different sets of measurements. The second round of measurements displayed much less variation than the first, although there was still more variation within some datasets than others. For example, the armspan measurements had more variation than did footlength. A class discussion on the differences between the first and second sets of data emphasized the possible sources of variation in each data set.

Assessment for this lesson occurred during the class discussion in which participants' responses revealed their understanding of variation within a dataset, reliability in data gathering, sources of variability, and measurement error.

This lesson's focus on the importance of considering the sources of variation within a dataset challenged participants to address the questions: Could some of the variation in a particular dataset have come from unreliable or sloppy data gathering techniques? Or, is all the variation inherent in the quantity being measured (that is, random fluctuation)? An important realization in this lesson is that there are multiple sources of variability, including variation inherent in a measured quantity.

The *Chip Sampling from Known and Unknown Mixtures* Lesson

Teachers and students are not typically asked to explore information from repeated sampling. Often just one sample is drawn from a population, and either the population information is used to predict what will be in that one sample, or the sample results are used to estimate population parameters. Thus, teachers and students rarely have an opportunity to explore the *variability* that occurs in sampling results when repeated samples are drawn. Research has revealed that some people believe there should be very little variability from sample to sample in re-sampling situations (Shaughnessy, J.M., 2006). For example, when given a mixture of yellow and green chips some students might predict that since the mixture is 70% yellow, then in samples of size 10 there *should* be 7 yellow every time. Of course since sampling is not "perfect," sometimes you'll get 6 or 8, or even 9 yellow. Some people predict a very "tight" spread of results in re-sampling. Others believe that just about anything can happen, and that results from re-sampling will be all spread out and that the range of possible outcomes from 0 yellow to 10 yellow in a sample of 10 chips will all occur. This tension, between representativeness (centers) and variability (spread) in sampling distributions is the centerpiece of the Chip Sampling lesson, enabling participants to consider the likely range of sample variation if repeated samples are drawn.

The focus questions for this lesson were, "If you knew the percentages of various objects in a very large mixture, what would you expect to see in a sample of those objects? Suppose you *didn't* have information about a very large mixture? How could you estimate what the mixture percentages are?" Objectives for the participants were:

- to introduce the concept of a sampling distribution and develop a feel for "likely range" and 'likely shape' of a binomial sampling distribution;
- to begin to make connections between (theoretical) probability distributions and sampling distributions, and begin to understand how sampling distributions can help to make decisions under uncertainty; and
- to introduce and highlight the importance of thinking in terms of 'likely intervals' for an outcome versus an exact point value.

In order to understand why population parameter estimates are often given in the form of confidence intervals, or with margins of error (e.g., 47% ± 2%), teachers and students need opportunities to create their own empirical sampling distributions, giving them experience with the likely range of samples. In the *Chip Sampling from Known and Unknown Mixtures* lesson, participants were able to draw repeated samples from both known and unknown mixtures, and create empirical sampling distributions of their results. For known mixtures, participants first created a predicted sampling distribution based on their knowledge of the mixture percentages, and then compared it to their actual results obtained by taking repeated samples from the mixture. For unknown mixtures, participants generated empirical sampling distributions and then used them to estimate mixture percentages.

This lesson also had the potential to lead to discussions of beliefs about probability, of expected results from sampling, and beliefs about the occurrence of outliers and how people expect extremes to occur more often than they actually ever do. The chip sampling activity addressed the core ideas of using data to make predictions and confidence in decision-making.

For the first task, each group was given a plastic container with 100 chips in it of two colors, in a 70-30 yellow-green mixture. Groups used grid poster paper to construct dot-plots of their predictions and their sampling results on the poster paper to post and share on walls of the classroom.

The participants were asked to predict the outcomes for the numbers of yellows in 30 samples of 10 chips. First they were given private think time to make their individual predictions. Next, groups discussed their individual predictions, came to group consensus, prepared and posted a dot-plot frequency distribution for their group predictions. In a whole class discussion, groups were asked what they noticed, and wondered about as they compared the various groups' predictions for the sampling task. Are there similarities or differences in the predictions? Why did groups make those particular predictions and what is the thinking behind the predictions?

After discussing the predicted sampling distributions, groups were instructed to draw 30 samples of 10 chips, with replacement between samples, from their containers and post their actual sampling distributions next to their group's predicted sampling distributions. Discussion of the comparison between the predicted distributions and the empirical distributions focused on what they noticed, what they wondered about, and some specific questions:

- What is a "likely range" for the number of yellows in a sample of 10 chips pulled from this mixture? (The idea here is that although there is variability in the results, most of the samples occur within a particular range, and so what would the likely range be for this mixture).
- How likely is it that "no yellows" would occur in a handful?
- Did it happen? Could it happen?
- How many samples do you think would need to be pulled to get at least one handful of 10 with no yellows in it?

Using a computer simulation of the sampling mixture (using either the Binomial Distribution Simulator Applet (Shaughnessy et al, 2004, CD), or other software such as *Tinkerplots* (Key Curriculum Press, 2005a) or *Fathom* (Key Curriculum Press, 2005b) that can generate empirical sampling distributions from mixtures) many, many samples were quickly drawn by the instructor while the participants watched as the distribution grew. Again a whole group discussion with participants focused on what they noticed, and what they wondered about. This activity provides the basis for future investigation of the theoretical binomial probability distribution for such sampling problems. In addition to investigating the sampling distribution for samples of 10 chips by using the computer to generate a large number of samples, the computer simulation could run many trials of 30 samples of 10, and count whether it ever happens that a "no yellow" sample occurs.

Having explored the nature of the sampling distribution for a known mixture (the 70-30 two color mix), the second task for each group was to determine the proportions in a two-color mixture of unknown composition. Here a much larger mixture is recommended, such as 1000 chips. One mixture that has proven interesting is a 55% - 45% mixture, as arguments will surface as to whether the mixture is, or is not, a 50-50 mixture. (This activity has a lot of room to be extended to hypothesis testing, and confidence interval comparisons, of various mixture possibilities, depending on the knowledge and background of the teachers, although we found that most of the high school teachers in our group were not ready for this extension.)

In the introduction of the unknown mixture task, we included a reminder that in the previous task everyone knew what the mixture was ahead of time, so they could make predictions based on the known population mixture. Most of the time, in real statistical sampling situations, we have no idea what the actual mixture in the population is. For the second task, the groups were given the large mixture with instructions to devise and carry out a sampling procedure (pull samples and remix each time, as in task one) to gather information that would help the group estimate the mixture percentages in the large container. For this task, the maximum allowable sample size was 20 chips. Each group decided on a sample size and the number of samples to draw, created a dot plot of the results of their samples (an empirical sampling distribution), and posted their results along with their group decision on the percentages of the mixture in the unknown container. Every member of each group was expected to be prepared to defend/explain the group's decision-making mechanism.

All groups quickly summarized their results and their decision for the other groups, being certain to justify their approach using statistically sound arguments for what they had done. A whole class discussion and question session followed that provided opportunities for groups to ask questions of other groups, compare and contrast these results, and see if the class could come to a consensus on estimating the mixture in the large containers. The discussion revealed the participants' desire to construct an interval estimate of the mixture, foreshadowing the upcoming lessons on confidence intervals.

As noted earlier, the main goal of this lesson was to develop participants' intuition for the likely range of variability in a sampling-resampling situation, when an empirical sampling distribution is created. Prompts such as these were helpful during discussions in focusing participants on this issue:

a) Recall our small container task, with a 70%-30% mixture. Suppose that we pull samples of 10 chips at a time. Where would you expect about 95% of the samples to fall, between _______ # yellows and ________ # yellows? Why would you predict that? Explain your reasoning.

b) Now suppose that a container had a 50% - 50% mixture of two colors. Suppose that we pull samples of 10 chips at a time. Where would you expect about 95% of these samples to fall, between _______ # yellows and ________ # yellows? Why would you predict that? Explain your reasoning.

Since the participants had just completed the sampling, formulated a decision, defended their decision, and asked questions of their classmates, these prompts invited them to generalize what they had learned and pushed them to justify their conclusions.

Experience with shape, center, and spread of sampling distributions is critical to this lesson. Therefore, as part of the "look back" piece in this lesson, participants were asked to reflect on what they had learned about sampling distributions. How did actual results compare to their predictions? Were they surprised—what did they learn, what would they predict now, after having conducted some repeated sampling experiments? How has their thinking changed about sampling? These questions were addressed in the whole group discussions.

Impact of the Professional Development Institute

The overarching focus on the use of justification and generalization in all Institute courses resulted in significant attitude and perspective changes on the part of the participants. Many of them were uncomfortable with these notions initially, but gradually came to understand the usefulness and power of this emphasis. Right answers alone were insufficient; process and explanations were stressed. Institute faculty continually pushed participants for justifications of their thinking. Whenever possible participants were encouraged to generalize which proved to be difficult for many teachers initially, but became somewhat easier by the third summer of the Institute.

Participants were required to write reflective journal entries during the content and leadership courses in an attempt to focus their attention on classroom discourse. Faculty were careful to model the type of discourse one might expect to find associated with a push for justification and generalization. Follow-up research was conducted both during the Institute and later in the teachers' own classrooms to document and categorize the types of discourse that were evident, and to determine the quantity and quality of high level discourse that one might associate with classrooms where justification and generalization were the norms.

It is not possible to assess the impact on teacher and student learning of any one of the OMLI courses in isolation. Mathematics content was coordinated, shared, and interwoven throughout the six courses. However, to measure overall gains in content knowledge, the project used instruments developed by the Study of Instructional Improvement and The Learning Mathematics for Teaching (LMT) Project at the University of Michigan (LMT, 2010). Content strands represented in the assessment items include number concepts, operations, patterns, functions, algebra, and geometry. Items on the instruments are designed to measure "mathematics content knowledge for teaching" in the sense of the mathematics that teachers use in the classroom daily to make instructional decisions (for example, in judging whether a student's alternative approach is mathematically sound). After the third OMLI summer institute, when teachers had completed all six courses, significant growth over their pre-test scores was evident on all subscales and overall on the Learning Mathematics for Teaching measures.

Upon completion of the Institute, teachers were expected to return to their schools and implement the lessons learned in their classrooms. It is important to note here that it was not the goal of OMLI to provide teachers with specific mathematics content and activities for their own classrooms. Rather, OMLI provided teachers with new mathematical practices they could use with their students to stimulate reasoning using generalizations and justifications in classroom discourse. It was this emphasis on mathematical practices that meant the course design concentrated on just a few big ideas as a modeling vehicle rather than attempting to cover a large and comprehensive list of mathematical topics that could be reproduced in the teachers' classrooms.

Lessons Learned

Others planning such a program to enhance teacher content knowledge should be aware that working with K-12 teachers in a heterogeneous classroom is both a challenging and an incredibly rich and rewarding experience. Teachers at all levels learned so much by working side by side and had the rare opportunity to consider the developmental "trajectory" of mathematical ideas across the K-12 span. At the same time, it is downright hard work. The faculty had to learn and hone skills regarding activity-based group learning, group protocols, social and socio-mathematical norms, status issues and mathematical discourse, and to make appropriate adjustments over the three years of the institute. They had to be prepared to deal with the inevitable status issues that arose, for example, when elementary teachers felt inadequate, or when high school teachers felt that they had nothing to learn. The first year of the Institute was the most challenging; the participating teachers had to learn to leave their egos and status issues by the door on the way in to each class, and to be open to learning from other teachers at all levels.

The Institute faculty had anticipated some of these issues, and had spent many hours collaborating on the design of the courses, selecting tasks that offered multiple access points for teachers at all levels. Teachers from each of the participating districts worked together over the three years of the Institute, and as a result teachers from different grades came to know and understand the work of their fellow teachers in earlier and subsequent grades; bonds and friendships were solidified. The fact that teachers themselves were empowered with the content knowledge and the leadership skills necessary for not only implementing productive mathematical practices in their own classrooms, but also to engage fellow teachers in professional development centered around student mathematical thinking. will help in sustaining OMLI's efforts.

In addition, we learned that it takes time to develop new skills and "habits of mind" in teachers. Working with teachers to include in their instruction a focus on generalization and justification was a significant challenge. It was essential to the success of the program that the teachers had three years of participation to learn about and practice these skills. Many teachers had not previously grappled with what it meant to generalize and justify mathematical ideas and needed reinforcement of and practice with the requisite skills.

The composition of the instructional teams – university faculty, community college faculty, and K-12 master teachers or mathematics specialists – was essential in the grounding of the works for K-12 teachers. Often, higher education faculty were not as familiar with working with teachers, especially at the elementary level, and had overly ambitious and unrealistic goals for the content courses. The master teachers frequently served to pull the team back to a "less breadth, more depth" approach, and to emphasize a focus on fundamental mathematical concepts. The instructional teams worked hard to make the connections to various points in the K-12 curriculum explicit. The course materials were not designed as "take back to your classroom" lessons. Rather, they were designed to provide as a vehicle for an in-depth look at fundamental concepts in the new light of justification and generalization. Many participants had never explored their curriculum from this perspective. By including teachers from kindergarten through high school, participants had the opportunity to embrace and grapple with the trajectory of a mathematical idea across the K-12 spectrum, and to appreciate the critical and essential nature of the focus on generalization and justification. Teachers were guided in learning how to incorporate these skills into their own classroom instruction, and to build a classroom expectation or norm for generalization and justification with their students.

Another lesson learned concerned the breadth of the courses and strategies for reducing the number of topics addressed in a given course while maintaining a coherent mathematical progression. As the OMLI faculty began implementing the courses, it became apparent that the initial plans for topics to address in the courses were too ambitious for the teacher participants. Thus, the plans for subsequent offerings of the courses were modified, using the focus on big ideas and coordination between the two courses in each pairing as principles for making content selection decisions. For example, the original design for Measurement and Change started with early measurement and change concepts and progressed up to calculus. As faculty reworked the course they focused on measurement and change in one dimension, in two dimensions, and in three dimensions as the organizing “big ideas.” Topics not central to those themes or some that were also addressed in Algebra, such as function were either dropped or deemphasized. Overall, the OMLI experience proved to be as profound a professional development activity for the institute faculty as it was for the participants.

References

American Statistical Association (2010). "What is a Statistical Poster?" The American Statistical Association Poster Competition. Retrieved from http://www.amstat.org/education/posterprojects/whatisastatposter.cfm.

Bright, G. W., Brewer, W., McClain, K., & Mooney, E. S. (2003). *Navigating through data analysis in grades 6-8.* Reston, VA: National Council of Teachers of Mathematics.

Bright, G. W., Frierson, D., Tarr, J., & Thomas, C. (2003). *Navigating through probability in grades 6-8.* Reston, VA: National Council of Teachers of Mathematics.

Burrill, G., Franklin, C., Godbold, L., & Young, L. (2004). *Navigating through data analysis in grades 9-12.* Reston, VA: National Council of Teachers of Mathematics.

Chapin, S., Koziol, A., MacPherson, J., & Rezba, C. (2002). *Navigating through data analysis and probability in grades 3-5.* Reston, VA: National Council of Teachers of Mathematics.

Cohen, E.G. (1994). *Designing Groupwork Strategies for Heterogeneous Classrooms, 2nd Edition*. New York: Teachers College Press.

Cohen, E.G., Lotan, R.A., Scarloss, B.A., & Arellano, A. R. (1999). Complex instruction: Equity in cooperative learning classrooms. *Theory into Practice*, *38*, 80-86.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Schaeffer, R. Guidelines for Assessment and Instruction in Statistics Education Report (GAISE). Retrieved from http://www.amstat.org/education/gaise/, January 17, 2010.

Friel, S. N., Bright, G, W, & Curcio, F. R. (1997). Understanding students' understanding of graphs. *Mathematics Teaching in the Middle School*, *3,* 224-227.

Key Curriculum Press. (2005a). Tinkerplots Dynamic Data Exploration (Version 1.0) [Computer software]. Emeryville, CA: Author.

Key Curriculum Press. (2005b). Fathom Dynamic Data Software (Version 2.0) [Computer software]. Emeryville, CA: Author.

Konold, C. (1996). Representing probabilities with pipe diagrams. *Mathematics Teacher, 89,* 378-382.

Learning Mathematics for Teaching Project. (2010). Retrieved January 18, 2010, from http://sitemaker.umich.edu/lmt/home

Matsumoto, A. (1981). Correlation junior varsity style. In Shulte, A. P. (Ed.), *Teaching Statistics and Probability*, 126-134. Reston, VA: NCTM.

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education*, *26,* 20-39.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for K-12 mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2006). *Thinking and reasoning with data and chance.* Reston, VA: Author.

Penner, E. & Lehrer, R. (2000). The shape of fairness. *Teaching Children Mathematics, 7,* 210-214.

Shaughnessy, J. M. (2006). Research on students' understanding of some big concepts in statistics. In G. E. Burrill & P. C. Elliott (eds.), *Thinking and reasoning with data and chance: Sixty-eighth yearbook* (pp.77 – 98). Reston, VA: NCTM.

Shaughnessy, M., & Arcidiacono, M. (1993). Activity 7: Monty's dilemma, *Math and the Mind's Eye Unit VIII: Visual Encounters with Chanc*e. Salem, OR: The Math Learning Center.

Shaughnessy, J. M., & Chance, B. (2005). *Statistical questions from the classroom.* Reston, VA: NCTM.

Shaughnessy, J. M., & Dick, T.P. (1991). Monty's Dilemma: Should you stick or switch? *The Mathematics Teacher*, *84*, 252-256.

Shaughnessy, J. M., & Pfannkuch, M. (2002). How faithful is Old Faithful? Statistical thinking: A story of variation and prediction. *The Mathematics Teacher, 95,* 252-259.

Shaughnessy, J. M., Barrett, G., Billstein, R., Kranendonk, H. A., & Peck, R. (2004). *Navigating through probability in grades 9-12.* Reston, VA: National Council of Teachers of Mathematics.

Sheffield, L. J., Cavanagh, M., Dacey, L., Findell, C. R., Greenes, C. E., & Small, M. (2002). *Navigating through data analysis and probability prekindergarten-grade 2.* Reston, VA: National Council of Teachers of Mathematics.

Takis, S. L. (1999). Titanic: A statistical exploration. *The Mathematics Teacher, 92,* 660-664.

Watson, J. M., & Shaughnessy, J. M. (2004). Proportional reasoning: Lessons from research in data and chance. *Mathematics Teaching in the Middle School, 10,* 104-109.

Weaver, D., & Dick, T. (2009). Oregon Mathematics Leadership Institute Project: Evaluation results on teacher content knowledge, implementation fidelity, and student achievement. *The Journal of Mathematics and Science: Collaborative Explorations*, *11*, 57 – 84.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*, 223-265.

Appendix A: Resources

Body Measurements Lesson

Supplies needed: Measuring tapes (60 inch/155cm or less), graph paper, butcher paper with scales ready to go for dot plots (2 sets), colored sticky dots for creating the repeated measurement graphs.

Classroom setup: Provide measuring tapes for each participant. These measuring tapes should be too short to measure armspan. Set up 4 measuring stations with:

(a) Poster graph for a dot plot of the measurements the participants will make. Be sure the horizontal axis can accommodate the expected range of measurements for the station attribute, *armspan, headspan, forearm length* or *footlength*. Prepare 2 of these charts in advance for each station.
(b) Sticky dots for recording measurements, and
(c) a chair or stool for the person being measured.

Sampling from Known and Unknown Mixtures Lesson

Supplies Needed: The binomial applet from the CD in the NCTM Navigating Probability book, 9-12 (or some other application that can generate graphs of results of repeated sampling in binomial trials, such as *Tinkerplots, Fathom*, etc). **Lots** of small plastic chips of two different colors, say yellow and green. For the first task, each group will need a container with a mixture of 100 chips (70 yellow and 30 green in the version written up here), and for the second task each group will need a mixture of 1000 chips (recommend 55% yellow, used in this lesson, as it leads to interesting discussion across groups when they compare results). Also recommended: **Lots** of dots that peel off and can be used to create dot plots of sampling results, and tear off grid poster paper to graph and share the results of empirical sampling distributions.

Body Measurements (Student page)

1. Measure (in centimeters) and record each of the four desired attributes:

Armspan	Headspan	Forearm	Foot length

   Also record your measurements on the class data tables and plot your values on the graph.

2. Examine the graphs of the data collected by the class. Record your observations about the body measurements taken by you and your classmates.

3. Using criteria established by the class for each measurement, measure and record each of the four desired attributes:

Armspan	Headspan	Forearm	Foot length

   Also record your measurements on the class data tables and plot your values on the graph.

4. Examine this new set of graphs and record your observations about the data.