RESEARCH K-12

Preservice Secondary Mathematics Teachers’ Statistical Knowledge: A Snapshot of Strengths and Weaknesses

Jennifer N. Lovett a and Hollylynne S. Lee b

a Mathematical Sciences, Middle Tennessee State University, Murfreesboro, TN; b North Carolina State University, Raleigh, NC

ABSTRACT

Amid the implementation of new curriculum standard regarding statistics and new recommendations for preservice secondary mathematics teachers (PSMTs) to teach statistics, there is a need to examine the current state of PSMTs’ knowledge of the statistical content they will be expected to teach. This study reports on the statistical knowledge of 217 PSMTs from a purposeful sample of 18 universities across the United States. The results show that PSMTs may not have strong Common Statistical Knowledge that is needed to teach statistics to high school students. PSMTs’ strengths include identifying appropriate measures of center, while weaknesses involve issues with variability, sampling distributions, p-values, and confidence intervals.

KEYWORDS

Education; Other; Testing

1. Introduction

Many have argued the need to increase students’ understanding of statistics (Shaughnessy 2007). Accordingly, there has been recent increased emphasis on statistics content in secondary curricula standards in the U.S., informed by recommendations from the National Council of Teachers of Mathematics (2000) and the Common Core Standards for Mathematics (CCSSM) (National Governors Association Center for Best Practice & Council of Chief State School Officers 2010). However, a recent study of 1,249 high school students in the U.S. suggests that students are not developing a conceptual understanding of statistics (Jacobbe, Foti, Case, and Whitaker 2014). Since many teachers, including preservice secondary mathematics teachers (PSMTs), have likely had minimal experience with statistics in their own K-12 education, they may not have had many opportunities to develop strong statistical understandings.

The Conference Board of the Mathematical Sciences (2001, 2012) as well as the American Statistical Association (ASA; Franklin et al. 2015), present recommendations for developing statistical knowledge and pedagogy needed by preservice secondary mathematics teachers (PSMTs) to teach statistics. However, the lack of research focusing on the statistical knowledge of PSMTs was highlighted and called for in the 2011 International Congress of Mathematics Education Topical Study (Batanero, Burrill, and Reading 2011). The majority of research on preservice teachers’ statistical knowledge has focused on elementary teachers (e.g., Browning, Goss, and Smith 2014; Groth and Bergner 2006; Hu 2015; Leavy 2010; Leavy and O’Loughlin 2006; Santos and da Ponte 2014). The limited research conducted on PSMTs’ statistical knowledge has been small-scale studies, from a small number of institutions on specific statistical content (e.g., Doerr and Jacob 2011; Lesser, Wagler, and Abormegah 2014). While some smaller studies have suggested that PSMTs may struggle with statistics (e.g., Casey and Wasserman 2015), there are no large-scale studies that describe the current state of new teachers’ statistical knowledge. This study examines the statistical knowledge of a large cross-institutional sample of PSMTs as they enter student teaching to answer the question: What are the strengths and weaknesses of PSMTs’ knowledge of the statistical content they will be expected to teach?

2. Background Literature

Building from the work of Hill, Ball, and Schilling (2008), Groth (2013) developed a hypothetical framework for Statistical Knowledge for Teaching. This framework consists of two domains of knowledge teachers need to develop: Subject Matter Knowledge and Pedagogical Content Knowledge. Developing Subject Matter Knowledge and Key Developmental Understandings of statistics is foundational to be able to develop Pedagogical Statistical Knowledge. Within subject matter knowledge there are three types: Common Statistical Knowledge, Specialized Statistical Knowledge and Horizon Knowledge. Common Statistical Knowledge refers to knowledge gained through statistics taught in school and is considered common because it refers to knowledge for daily literacy or in any profession that uses statistics. This study focuses on the common statistical knowledge of PSMTs since it is foundational for developing Pedagogical Statistical Knowledge. Common Statistical Knowledge is also the content that they will soon be expected to teach as part of curricula for high school students.
The *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework* (GAISE; Franklin et al. 2007) describes statistical reasoning students should develop in K-12 and thus setting a minimum Common Statistical Knowledge that PSMTs should pose to be prepared to teach statistics. The GAISE framework consists of three levels A, B, and C across which statistical reasoning develops. Even though GAISE does not provide explicit definitions given for each level, as the levels increase A to C they become more abstract and increase in sophistication. The content in Level A represents topics for early or novice learners of statistics, Level B represents slightly more advanced statistical content, and Level C represents even more advanced content (Franklin et al. 2007). The authors of the report hope that content at level A is given for each level, as the levels increase A to C they become prepared to teach statistics. The GAISE framework consists of Common Statistical Knowledge that PSMTs should pose to be prepared to teach statistics. The GAISE report recommends that students learn statistical topics through engaging in a statistical investigative cycle consisting of: posing questions, collecting data, analyzing data, and interpreting results. Therefore, when examining PSMTs’ Common Statistical Knowledge, it is useful to consider their understandings across these cycle phases and all three GAISE levels.

### 2.1. PSMTs’ Statistical Knowledge

As previously mentioned the lack of research focusing on PSMT’s and teachers’ statistical knowledge was highlighted in the 2011 International Congress of Mathematics Educational Topical Study (Batanero et al. 2011). From the limited research conducted on PSMT’s and teachers’ statistical knowledge several themes have emerged: (1) focus on procedures, computations, and algorithms, (2) lack of reasoning skills, and (3) difficulty constructing and interpreting graphical representations.

The focus on procedures, computations and formulas has been demonstrated through several studies. In regards to the mean, teachers incorporate two approaches: balancing point and standard algorithm but often rely on the formula to define and calculate the mean (Gfeller, Niess, and Lederman 1999; Russell and Mokros 1990). Research also shows that teachers are fluent with the procedures of descriptive statistics (Makar and Confrey 2004) and in turn concentrate on computations and creating graphical displays and lose focus on the statistical investigation (Burgess 2002; Heaton and Mickelson 2002). The fact that this theme has emerged in the research is not surprising. Mathematics often has a correct answer and statistical reasoning requires drawing conclusions that are uncertain (Groth and Bergner 2007; Rossmann, Chance, and Medina 2006).

Since teachers often focus on the computations and creating graphical representations, researchers would expect that teachers would excel in the construction and interpretation of graphical representations. However Meletiou-Mavrotheris and Lee (2003) found that teachers struggled with the construction and interpretation of graphical representations. Specifically, teachers in their study struggled constructing histograms and box plots, interpreting histograms and box plots, and confused a bar graph with a histogram.

This concentration on computation may cause teachers to struggle with statistical reasoning skills identified in the research. One area that has been studied by several researchers is teachers’ abilities to interpret and analyze distributions. Researchers have found that teachers are confident in analyzing and interpreting symmetric distributions but struggle with skewed distributions (Doerr and Jacob 2011; Rubin, Hammerman, Campbell, and Puttik 2005). One example is the research Doerr and Jacob (2011) conducted with secondary mathematics teachers after taking a one-semester course that engaged teachers in statistical investigations. On a post-test they found that 18% of the teachers did not understand that a distribution with the median larger than the mean is likely to be skewed left (Doerr and Jacob 2011). Teachers also encounter difficulty distinguishing the sampling distribution from the population distribution. Researchers found teachers struggled differentiating between the variability of the distribution from the sampling distribution (Makar and Confrey 2004) and describing the sampling distribution to be the same as the population distribution (Doerr and Jacob 2011).

To date, only one large-scale study, that we know of, was conducted by Lee et al. (2014). They examined how 204 preservice mathematics teachers from eight universities used dynamic statistical tools to conduct a statistical investigation. They found that preservice teachers who pose a broad statistical question engaged in more graphical augmentations (e.g., adding shaded regions, reference lines, or statistical measures) using dynamic statistical software. These graphical augmentations allowed preservice teachers to dive deeper into the data analysis and make connections to the context to support claims. For the field to truly understand PSMTs’ statistical knowledge, 2011 International Congress of Mathematics Educational Topical Study called for more studies to be conducted (Batanero et al. 2011).

### 3. Methodology

The results reported here are part of a larger mixed methods study on preparedness of PSMTs to teach statistics (Lovett 2016). While the larger study included instruments to measure self-efficacy to teach statistics and interviews with some PSMTs about experiences in their teacher preparation programs, we only focus here on quantitative analysis of a measure of their Common Statistical Knowledge.

#### 3.1. Participating Institutions

This study focuses on PSMTs prepared through university-based teacher preparation programs in the US. Since a random sample of all mathematics teacher preparation programs was unavailable, this study began with a purposeful narrowing on two nationally-funded programs that included considerable effort to assist mathematics teacher educators and/or mathematics and statistics faculty to develop new understandings and approaches to preparing teachers to teach statistics. There were 57 different institutions from which some faculty members
participated in the NSF-funded program, Preparing to Teach Mathematics with Technology (PTMT, ptmt.fi.ncsu.edu), and/or the ASA-funded Math/Stat Teacher Education: Assessment, Methods, and Strategies (TEAMS, www.amstat.org/sections/educ/newsletter/v9n1/TEAMS.html) conference between 2002–2014. These institutions were chosen since faculty members received professional development focusing on explicit content and strategies for preparing PSMTs to teach statistics. Our assumption is that faculty from these institutions may have been motivated to include more of a focus on statistical content. Thus, PSMTs from these institutions represent a critical case (Patton 2002) since, perhaps, they have had more opportunities in their coursework than may be represented at other institutions.

The sample was obtained by contacting all 57 institutions through their undergraduate program coordinator for mathematics education to inquire if the program was interested in participating. Twenty-four programs expressed interest, and 18 participated. The coordinator identified the last mathematics teaching methods course PSMTs take before student teaching, which would constitute the data collection point in either Fall 2014 or Spring 2015. Of the 18 institutions, all but one were public institutions. The Carnegie Classification™ (Carnegie Foundation for the Advancement of Teaching 2011) for institutions participating in the study is displayed in Table 1.

### 3.2. Participants

Across 18 institutions, there were 221 PSMTs recruited by their mathematics teaching methods instructor to take the assessment of their Common Statistical Knowledge, described in the next section, as an assignment as part of the course. Those who took exceptionally less time (10 minutes) than recommended by authors of the assessment were eliminated (Jacobbe, personal communication, August 11, 2015). This resulted in a sample size of 217 PSMTs. The PSMTs were undergraduate juniors and seniors, or graduate students earning initial licensure; all were enrolled in their last mathematics education course prior to student teaching. The number of PSMTs participating from each institution ranged from 2 to 31, with a mean of 12. Fourteen institutions had 100% participation of PSMTs who were eligible to participate, with the remaining four institutions having between one and four students who did not complete the assignment. The majority of PSMTs were female (71%), and 88% were Caucasian. Almost all (93.4%) reported they had taken at least one statistics course at their institution or had completed Advanced Placement Statistics in high school. The demographics of the PSMTs in regards to taking AP Statistics in high school, the self-reported number of college-level statistics courses taken, and the degree program in which they were enrolled are shown in Table 2.

#### Table 2. Demographics of the 217 participating PSMTs.

| Characteristic of PSMTs | n (%) |
|-------------------------|-------|
| Gender                  |       |
| Female                  | 154 (71) |
| Male                    | 63 (29) |
| High School Course      |       |
| AP Statistics           | 39 (18) |
| Number of University Statistics Courses |
| 0                       | 10 (4.6) |
| 1–2                     | 128 (59) |
| 3 or more               | 79 (36.4) |
| Degree Program          |       |
| Bachelor’s              | 201 (92.6) |
| Master’s                | 16 (7.4) |
| Bachelor’s Diverse      | 109 (50.2) |
| Master’s Diverse        | 17 (7.9) |
| Bachelor’s Smaller      | 82 (38.2) |
| Master’s Smaller        | 8 (3.7) |

#### Table 1. Carnegie Classification™ of participating institutions.

| Enrollment Profile | Universities n (%) | Participants n (%) |
|--------------------|--------------------|--------------------|
| VHU: Very High Undergraduate | 7 (38.9) | 68 (31.3) |
| HU: High Undergraduate | 11 (61.1) | 149 (68.7) |
| Basic              |        |        |
| RU/VH: Research Universities (Very High) | 4 (22.2) | 67 (30.9) |
| RU/H: Research Universities (High) | 6 (33.3) | 54 (24.9) |
| D/RU: Doctoral/Research Universities | 2 (11.1) | 16 (7.4) |
| Master’s: Master’s Colleges and Universities (Larger) | 4 (22.2) | 61 (28.1) |
| Master’s: Master’s Colleges and Universities (Smaller) | 1 (5.6) | 6 (2.8) |
| Bac/Diverse: Baccalaureate College – Diverse Fields | 1 (5.6) | 13 (5.9) |

### 3.3. Instrument

To examine PSMTs’ Common Statistical Knowledge, the Levels of Conceptual Understanding of Statistics (LOCUS) assessment (Jacobbe, Case, Whitaker, and Foti 2014) was administered online (locus.statisticseducation.org). This instrument was developed using an evidence-centered design approach through content domain analysis using the GAISE framework, the statistics content in the CCSSM, and learning trajectories from research on students’ learning in statistics (Haberstroh et al. 2015). The LOCUS instrument is aligned with and assesses statistical content across the three GAISE levels and within each phase of an investigative cycle: formulate questions, collect data, analyze data, and interpret results.

The instrument authors’ descriptions of items from each of the four phases of the investigative cycle are displayed in Table 3 (Jacobbe, Foti, et al. 2014). Actual test items cannot be released due to test security; sample items for the four categories at different levels are available on the LOCUS website (locus.statisticseducation.org/professional-development) and sample items are in the results section.

Participants took the 30 multiple choice Intermediate/Advanced Statistical Literacy version of the assessment, which was designed for students in grades 10–12. The test consists of two level A questions, 11 level B questions, and 17 level C questions. This version has been validated and reliable with students in grades 6–12 to assess statistical knowledge across levels B and C and the four phases of the investigative cycle (Jacobbe, personal communication, June 7, 2016); while this instrument is not intended as a high stakes assessment of knowledge, it does represent the statistics content PSMTs are expected to teach their students in the near future. Thus, teachers are expected to score fairly high on the assessment. Each test-taker receives an overall score (percent correct), as well as sub-scores for Level B, Level C, Formulating Questions, Collecting Data, Analyzing Data, and Interpreting Results.
Table 3. Description of LOCUS items.

| Phase                  | Description                                                                 |
|------------------------|-----------------------------------------------------------------------------|
| Formulate Questions    | Involve the planning process used in a statistical problem, including having the students determine if the problem is statistical in nature |
| Collect Data           | Involve executing or implementing sampling or experimental assignment of treatment techniques |
| Analyze Data           | Involve students analyze data through a statistical lens to show that they understand what the data is telling them. |
| Interpret Data         | Involve the students answering an initial question by drawing conclusions from the data. |

3.4. Data Analysis

To examine the Common Statistical Knowledge demonstrated by PSMTs, descriptive statistics were computed for the overall score and each subscore. A quantitative item analysis was conducted by examining the performance of each individual item across all PSMTs to closely examine PSMTs' strengths and weaknesses. In such an item analysis, we were able to look more carefully at the mean percent of all PSMTs that chose the correct response, as well as examine patterns in selection of incorrect responses.

4. Results

To situate the results of the item analysis, first descriptive results are presented of PSMTs’ LOCUS overall scores and subscores across GAISE levels and phases of the statistical investigative cycle. Then an item analysis is presented with results organized by the alignment of questions to the statistical investigative cycle.

4.1. Overall Scores

Trends in scores on the LOCUS test can help in describing what PSMTs from these 18 universities currently understand about the statistics content they will soon be responsible for teaching. PSMTs’ scores had a mean overall score of 69%, and a standard deviation of 14.06. Since this is an assessment designed for students, a mean score of 69% demonstrates that PSTs do not have a deep understanding of the statistical content they will teach. However, there is a concern since only one-quarter of PSMTs scored overall above 77%, and a quarter scored below 57% overall. The variation in scores seems somewhat similar for Level C scores. However, higher variability in Level B scores is due to the increased quantity of low scoring individuals.

Examining subscores by phases in the statistical investigative cycle, the data shows that for all four phases, there are again some PSMTs who scored between 90–100% on individual phases (Table 4). This indicates that those PSMTs likely have the Common Statistical Knowledge that will be needed when teaching that phase of the investigative cycle. On Formulating Questions items, at least half of PSMTs scored 80% or higher, and a quarter of those scored 100%, indicating stronger understanding for these PSMTs about Formulating Questions. However, half of PSMTs scored below 71% on Collecting Data and Analyzing Data items, and half scored below 64% on Interpreting Results items. Interestingly, only two PSMTs scored over 80% on all four phases. Even being conservative, this result is convincing that the vast majority of these PSMTs do not have the Common Statistical Knowledge that can provide a foundation for teaching students key concepts related to Collecting Data, Analyzing Data, and Interpreting Results. For an in depth analysis of PSTs’ LOCUS scores and factors that impacted these scores, see Lovett (2016).

4.2. Item Analysis

To identify PSMTs’ strengthens and weaknesses within each phase of the statistical investigative cycle, an item analysis of PSMTs’ performance on individual items was conducted.

4.2.1. Formulating Questions

As previously mentioned, PSMTs scored the highest, on average, on Formulating Questions items. The item difficulty for each item categorized by the authors of the LOCUS as formulating questions are displayed in Table 5. Even though four out of the five Formulating Questions items were written at a GAISE Level C, the majority of PSMTs answered each Formulating Question item correctly. In these questions, PSMTs were asked to read a description of a study and the measurements taken, to identify an appropriate statistical question of interest. An example of this type of item from the LOCUS professional development resources is shown in Figure 1. Item 12 posed the most difficulty for PSMTs of those that were categorized as Formulating Questions. This question asked PSMTs to identify which statistical question would be better answered by a sample rather than a census. So even though the majority of students answered this correctly, 39 percent of the PSMTs in this study have a weakness
in determining when a sample is appropriate to be used instead of a census.

### 4.2.2. Collecting Data

On average, PSMTs scored the next highest on Collecting Data items. The item difficulty for items categorized as Collecting Data are displayed in Table 6. Two of these seven items were aligned with GAISE Level B and the remaining items aligned with GAISE Level C. An examination of these items showed that PSMTs were able to identify ways to improve a study design given a study and measurements, identify which study design would be best based on a question of interest, and identify a data collection plan based on a study description. Thus, these PSMTs seem to have strong Common Statistical Knowledge related to the design of a statistical study.

Even though PSMTs were able to develop a data collection plan, they struggled more with item 6, when asked to identify how to choose a sample to minimize bias. Only 64.5% were able to choose a correct sampling method; instead, 30% chose a convenience sample or a stratified sample that seemed complicated but was not random. Thus, these PSMTs do not seem to have a strong understanding of the role of a random sampling method within the design of a study.

A weakness demonstrated by the majority of PSMTs was the conclusion that could be drawn from a specific study design. Figure 2 is a similar item to the one PSMTs were asked on the instrument. Over 58% of PSMTs chose an answer similar to answers (A) and (C) that allowed a researcher to generalize results to an entire population based on a sample of volunteers. These findings highlight PSMTs’ need for a deeper understanding related to ways in which study designs and data collection processes impact the conclusions that can be drawn.

### 4.2.3. Analyzing Data

PSMTs’ average scores for Analyzing Data items were the second lowest among the phases, and had the highest variability even though only two out of the seven items categorized as Analyzing Data items were at GAISE Level C. Table 7 displays the item difficulty for items categorized as Analyzing Data items. The majority of PSMTs demonstrated that they understand which measure of center is appropriate for a given context, how measures of center and variation change when data values are changed, how to read a box plot, and a justification of an association from a two-way table.

A common misunderstanding of PSMTs throughout the Analyzing Data items involved the understanding of variation in data. This was first demonstrated in item 16. The majority of PSMTs answered item 16 correctly, but 36% of PSMTs demonstrated a misunderstanding related to expected variation in sample means when repeatedly sampling from a population. When given the distribution of a population and population mean, 36% of PSMTs could not identify the distribution of sample means. Instead they chose distributions that resembled the general shape of population distribution. The other misunderstandings related to variability were demonstrated by the majority of PSMTs. For item 18 only 43% of PSMTs could identify a histogram containing data that varied the least from its mean. Instead 30% of PSMTs chose a uniform distribution and about 20% thought variability from the mean was the same for all three distributions. A similar item to the actual item regarding variation is shown in Figure 3. However, in this sample item they are asked to choose the distribution with data that varies most from the mean. These results point to PSMTs’ need for more Common Statistical Knowledge in regards to variation, sample distributions, and distribution of sample statistics.

### 4.2.4. Interpret Results

PSMTs scored the lowest on average on Interpret Results items. However, Table 8 shows that on five of the eleven Interpret Results items, 84% or more of PSMTs answered the items correctly. Four of these five are GAISE Level B items. PSMTs demonstrated they were able to compare distributions in a

| Table 6. Item difficulty of Collect Data items. |
|-----------------|-------|-----|
| Item Number     | Mean  | GAISE Level |
| 3               | 0.83  | C    |
| 6               | 0.65  | B    |
| 10              | 0.72  | B    |
| 13              | 0.73  | C    |
| 21              | 0.31  | C    |
| 25              | 0.89  | C    |
| 27              | 0.78  | C    |

NOTE: Mean is the proportion of PSMTs who correctly answered each item

| Table 7. Item difficulty of Analyzing Data items. |
|-----------------|-------|-----|
| Item Number     | Mean  | GAISE Level |
| 7               | 0.70  | B    |
| 11              | 0.77  | C    |
| 15              | 0.68  | A    |
| 16              | 0.64  | C    |
| 18              | 0.43  | B    |
| 20              | 0.53  | A    |
| 28              | 0.66  | B    |

NOTE: Mean is the proportion of PSMTs who correctly answered each item

Each member of a random sample of 1,000 adult males from the United States was asked a number of questions, including questions about height and annual income. When the responses were analyzed, it was determined that taller men had greater incomes than shorter men, on average, and the difference was statistically significant. Which of the following conclusions would be most appropriate based on these results?

(A) The study establishes that being tall causes men to have greater incomes, on average.
(B) The study establishes that being tall causes men to have greater incomes, on average, but this result only applies to the men in the sample.
(C) The study establishes that being tall causes men to have greater incomes, on average, than shorter men, but this result only applies to the men in the sample.
context using the center and spread, demonstrate an understanding of the effect of sample size on a sample mean, identify bias in sampling techniques, and interpret survey results with a given margin of error. These are important concepts often taught in middle and high school curricula.

On the other six Interpret Results items, the percentage of PSMTs responding correctly to these items ranged from 21% to 48%, and their misunderstandings were related to ideas of formal inference. These items asked students to reason about identifying and interpreting a p-value, interpreting a sampling distribution from a simulation, the effect of sample size on sampling distributions of sample means and explaining confidence intervals.

Several items were related to the ideas of p-value and statistical significance. About half (48%) of PSMTs were able to correctly answer one of the formal inference problems, which asked PSMTs to interpret results given a large \( p \)-value. Approximately 40% of PSMTs chose an answer that a large \( p \)-value allowed for the researcher to reject the null hypothesis and conclude the alternative hypothesis to be true. This misunderstanding of \( p \)-value increased on another item that instead of being given a \( p \)-value, PSMTs were asked to reason if a \( p \)-value would be large or small for comparing means of two distributions given data on a dotplot. Only 35% of PSMTs were able to correctly identify that the \( p \)-value would be small due to the large gap between distributions. Almost 47% incorrectly answered that the \( p \)-value would be large due to a large gap between the distributions. Another item relating to statistical significance that only 32% of PSMTs were able to correctly answered involved interpreting a sampling distribution from a simulation. These findings demonstrate that PSMTs on average do not have an understanding of what it means to be statistically significant and what a \( p \)-value represents, aspects of Common Statistical Knowledge, and included in many high school curricula that include simulation-based inference.

PSMTs also struggled with the effect of sample size on sampling distributions of sample means. PSMTs were given a dotplot of 100 sample means with a sample size of 45 and asked to identify the dotplot from the same population of 100 sample means with a sample size of 90. Almost 30% of PSMTs chose a dotplot similar to the dotplot with a sample size of 45,
demonstrating they did not understand the effect sample size would have on the distribution. The item PSMTs had the most difficulty with in Interpreting Results asked the test taker to explain the meaning of a 95% confidence interval for a mean. Approximately one-fifth chose the correct response that a 95% confidence interval represents that 95% of confidence intervals constructed from random samples would capture the true mean. Almost half of PSMTs chose the response that there was a 95% probability that the mean was in between the lower and upper limits of the confidence interval. These misunderstandings highlight the need for PSMTs to have more experiences with interpreting and understanding sampling distributions and confidence intervals.

5. Discussion

Our study was situated within a purposeful sample of PSMTs enrolled in teacher education programs where a faculty member had participated in professional development projects that promoted increasing attention to statistics in secondary mathematics education courses. It is not known exactly how those teacher education programs currently include an emphasis on statistics, nor exactly what these PSMTs experienced at all 18 institutions. Nonetheless, there are several findings of this study that are significant to consider.

Our results provide empirical evidence that approximately 25% of PSMTs in this study generally exhibit a strong Common Statistical Knowledge needed for teaching high school students. PSMTs’ knowledge was weaker in later phases of a statistical investigation. Previous research has shown a similar trend with inservice teachers and students measured by LOCUS (Jacobbe 2015; Jacobbe, Foti, et al. 2014). Thus, PSMTs need more experiences or different experiences in collecting data, analyzing data, and interpreting results to develop a deeper understanding of all aspects of the statistical investigative cycle and to develop Common Statistical Knowledge needed for teaching.

PSMTs exhibited some similar strengths and weaknesses with concepts that previous research has shown that high school and introductory college students develop. An important strength that PSMTs demonstrated is that they are proficient at identifying an appropriate measure of center for a given context. This topic is heavily emphasized in school mathematics, yet research shows that students struggle understanding measures of center (e.g., Mokros 1995; Watson 2007). PSMTs’ strength in understanding measures of center suggests they should be well-equipped to assist their future students develop stronger conceptions.

PSMTs’ weaknesses involve issues with variability, sampling distributions, p-values, and confidence intervals. Many researchers have identified that these topics are also often misunderstood by many students in undergraduate statistics courses (e.g., Aquilonius and Brenner 2015; Castro Sotos, Vanhoof, Van de Noortgate, and Onghena 2007; delMas, Garfield, Ooms, and Chance 2007); thus, PSMTs’ Common Statistical Knowledge may be no better than those of other college students not preparing for teaching.

All of the 18 universities included in this study require their PSMTs to take a statistics course and list some or all of these topics in their course descriptions. A sample of 25 PSMTs were interviewed about their experiences in these courses and they reported that the courses were typically lecture courses that were procedurally focused (Lovett and Lee 2017). However, we were not able to determine what statistical topics were taught in the courses. A challenge of conducting cross-institutional research is that individual programs and courses at each university are unique. Also, within a university a PSMTs’ experiences could vary due to instructor. Thus, we argue, even though emphasis on topics such as informal and formal inference should be occurring more in high school and college curricula, it appears that current experiences and opportunities for the PSMTs in this study have not been enough to develop a strong Common Statistical Knowledge of these topics. Their current experiences have not prepared PSMTs well to develop statistical knowledge for teaching (Groth 2013) and successfully teach high school students in ways that can develop their understanding of these statistics topics.

This research highlights the need for mathematics teacher education faculty and statistics faculty to work together to reexamine the statistics courses PSMTs take in their teacher education programs. In 2015, the ASA published The Statistical Education of Teachers (SET) report (Franklin et al. 2015). The SET report provides recommendations for statistics courses that should be taken by preservice secondary mathematics teachers (PSMTs) and the content of those courses. PSMTs’ weaknesses identified in this study are all topics that align with the topics that the SET report recommends to be emphasized in an introductory course (i.e., data analysis, simulation-based approaches to inference, and an introduction to formal inference). An introductory course would provide PSMTs with foundational statistical knowledge and experiences that they have not experienced in their own K-12 education and if not implementing the GAISE College Standards (2016) could help PSMTs develop an understanding of statistics beyond procedures.

6. Conclusion

These findings, even though from a purposeful sample, suggest there is a critical need for mathematics teacher education programs to reevaluate the opportunities PSMTs have to increase their Common Statistical Knowledge. Our results specifically indicate that effort should focus on developing PSMTs’ knowledge of variability, sampling distributions, and formal inference, particularly as they are applied in the analyzing data and interpreting results phases of an investigative cycle. More qualitative research is needed on how PSMTs’ knowledge of these areas develop to help statistics instructors and mathematics educators to design curricular experiences to strengthen this knowledge. Additionally, large-scale studies are needed on all aspects of PSMTs’ statistical knowledge for teaching and the impact that teacher education programs have on PSMTs’ preparedness to teach statistics. While our study did not examine the individual institutional programs, there is a need to further examine how mathematics teacher education programs are preparing PSMTs to teach statistics. Such case studies could provide recommendations to make large-scale changes for statistics courses and mathematics teacher education programs across the country.
Funding

The work in this study was partially supported by the Preparing to Teach Mathematics with Technology project, funded by the National Science Foundation with grants to NC State University (DUE 0442319, DUE 0817253, and DUE 1123001). The opinions, findings, and conclusions or recommendations are our own, and do not necessarily reflect the views of the National Science Foundation.

References

Aquilonius, B. C., and Brenner, M. E. (2015), “Students’ Reasoning about p-values,” Statistics Education Research Journal, 14(2), 7–27.

Batanero, C., Burrill, G., and Reading, C. (Eds.), (2011), Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE study, The Netherlands: Springer.

Browning, C., Goss, J., and Smith, D. (2014), “Statistical Knowledge for Teaching: Elementary Preservice Teachers,” in K. Makar, B. de Sousa, and R. Gould (Eds.), Sustainability in Statistics Education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July 2014) Flagstaff, Arizona, USA, Voorburg, The Netherlands: International Statistical Institute.

Burgess, T. (2002). “Investigating the ‘Data Sense’ of Preservice Teachers,” Paper presented at the Sixth International Conference on Teaching of Statistics, Cape Town.

Carnegie Foundation for the Advancement of Teaching (2011), The Carnegie Classification of Institutions of Higher Education, 2010 Edition, Menlo Park, CA: Author.

Casey, S. A., and Wasserman, N. H. (2015), “ Teachers’ Knowledge About Informal Line of Best Fit,” Statistics Education Research Journal, 14(1), 8–35.

Castro Sotos, A. E., Vanhooft, S., Van de Noortgate, W., and Onghena, P. (2007), “Students’ Misconceptions of Statistical Inference: A Review of the Empirical Evidence from Research on Statistics Education,” Educational Research Review, 2(2), 98–113.

Conference Board of the Mathematical Sciences (2001), The Mathematical Education of Teachers, Providence, RI: American Mathematical Society.

delMas, R., Garfield, J., Ooms, A., and Chance, B. (2007), “Assessing Students’ Conceptual Understanding After A First Course in Statistics,” Statistics Education Research Journal, 6(2), 28–58.

Doerr, H., and Jacob, B. (2011), “Investigating Secondary Teachers’ Statistical Understandings,” in M. Pytlak, T. Rowland, and E. Sobořá (Eds.), Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education (pp. 776–786).

Franklin, C., Bargagliotti, A. E., Case, C. A., Kader, G. D., Schaeffer, R. L., and Spangler, D. A. (2015), The Statistical Education of Teachers, American Statistical Association.

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

GAISE College Report ASA Revision Committee (2016), “Guidelines for Assessment and Instruction in Statistics Education College Report 2016,” Retrieved from http://www.amstat.org/education/gaisereport.

Gfeller, M. K., Niess, M. L., and Lederman, N. G. (1999), “Preservice Teachers’ Use of Multiple Representations in Solving Articmetric Mean Problems,” School Science and Mathematics, 99(5), 250–257.

Groth, R. E. (2013), “Characterizing Key Developmental Understandings and Pedagogically Powerful Ideas Within A Statistical Knowledge for Teaching Framework,” Mathematical Thinking and Learning, 15(2), 121–145.

Groth, R. E., and Bergner, J. A. (2006), “Preservice Elementary Teachers’ Conceptual and Procedural Knowledge of Mean, Median and Mode,” Mathematical Thinking and Learning, 8(1), 37–63.

Groth, R. E., and Bergner, J. A. (2007), “Teachers’ Perspectives on Mathematics Education Research Reports,” Teaching and Teacher Education, 23, 809–825.

Haberstroh, J., Jacobbe, T., delMas, R., Hartlaub, B., Miller, D., Schaeffer, R., ... Watson, J. (2015), “LOCUS Evidence Model,” retrieved from https://locus.statisticseducation.org/about-locus/evidence-model

Heaton, R. M., and Mickelson, W. T. (2002), “The Learning and Teaching of Statistical Investigation in Teaching and Teacher Education,” Journal of Mathematics Teacher Education, 5(1), 35–59.

Hill, H. C., Ball, D. L., and Schilling, S. G. (2008), “Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers’ Topic-Specific Knowledge Of Students,” Journal for Research in Mathematics Education, 39(4), 372–400.

Hu, S. (2015), “Pre-Service Teachers’ Understanding of Association and Correlation in An Inquiry-Based Introductory Statistics Course,” paper presented at the Thirty-seventh annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, East Lansing, MI: Michigan State University.

Jacobbe, T. (2015), “Developing K-12 Teachers’ Understanding of Statistics,” retrieved from https://www.causeweb.org/webinar/teaching/2015/04/7.

Jacobbe, T., Case, C., Whitaker, D., and Foti, S. (2014), “Establishing the Validity of the LOCUS Assessments Through an Evidenced-Centered Design Approach,” in K. Makar, B. de Sousa, and R. Gould (Eds.), Sustainability in Statistics Education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July 2014) Flagstaff, Arizona, USA, Voorburg, The Netherlands: International Statistical Institute.

Jacobbe, T., Foti, S., Case, C., and Whitaker, D. (2014), “High School (ages 14–18) Students’ Understanding of Statistics,” in K. Makar, B. de Sousa, and R. Gould (Eds.), Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July 2014) Flagstaff, Arizona, USA, Voorburg, The Netherlands: International Statistical Institute.

Leavy, A. M. (2010), “The Challenge of Preparing Preservice Teachers to Teach Informal Inferential Reasoning,” Statistics Education Research Journal, 9(1), 46–67.

Leavy, A. M., and O’Loughlin, N. (2006), “Preservice Teachers Understanding of The Mean: Moving Beyond the Arithmetic Average,” Journal of Mathematics Teacher Education, 9(1), 53–90.

Lee, H. S., Kersaint, G., Driskell, S., Harper, S. R., Jones, D., Leatham, K., ... Kwaku, A.-G. (2014), “Teachers’ Use of Translating in Solving Statistical Tasks With Dynamic Statistical Software,” Statistics Education Research Journal, 13(1), 25–52.

Lesser, L. M., Wagler, A. E., and Abornegah, P. (2014), “Finding a Happy Median: Another Balance Representation for Measures of Center,” Journal of Statistics Education, 22(3).

Lovett, J. N. (2016), The Preparedness of Preservice Secondary Mathematicians to Teach Statistics: A Cross-Institutional Mixed Methods Study, unpublished doctoral dissertation. Raleigh, NC: North Carolina State University.

Lovett, J. N., and Lee, H. S. (2017), “New Standards Require Teaching More Statistics: Are Preservice Secondary Mathematics Teachers Ready?,” Journal of Teacher Education, 68(3), 299–311.

Makar, K., and Confrey, J. (2004), “Secondary Teachers’ Reasoning About Comparing Two Groups,” in D. Ben-Zvi and J. Garfield (Eds.), The Challenge of Developing Statistical Literacy, Reasoning, and Thinking. Dordrecht, The Netherlands: Kluwer Academic Publisher.

Meletiou-Mavrotheris, M., and Lee, C. (2003), “Teaching Students the Stochastic Nature of Statistical Concepts in an Introductory Statistics Course,” Statistics Education Research Journal, 1(2), 22–37.

Mokros, J., and Russell, S. J. (1995), “Children’s Concepts of Average and Representativeness,” journal for research in Mathematics Education, 26(1), 20–39.

National Council of Teachers of Mathematics (2000), Principles and Standards for School Mathematics, Reston, VA: Author.

National Governors Association Center for Best Practice & Council of Chief State School Officers (2010), Common Core State Standards for Mathematics, Washington D.C.: Author.

Patton, M. Q. (2002), Qualitative Research & Evaluation Methods (3rd ed.). Thousand Oaks, CA: SAGE.

Rossman, A., Chance, B., and Medina, E. (2006), “Some Important Comparisons Between Statistics and Mathematics and Why Teachers Should Care,” in G. Burrill and P. C. Elliot (Eds.), Thinking and Reasoning with
Rubin, A., Hammerman, J. K., Campbell, C., and Puttick, G. (2005), “The Effect of Distributional Shape on Group Comparison Strategies,” in K. Makar (Ed.), Reasoning About Distribution: A Collection of Current Research Studies (SRTL-4), Brisbane: University of Queensland.

Russell, S. J., and Mokros, J. (1990), “What’s Typical: Teachers’ Description of Data,” paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.

Santos, R., and da Ponte, J. P. (2014), “Learning and Teaching Statistical Investigations: A Case Study of Prospective Teacher,” in K. Makar, B. de Sousa, and R. Gould (Eds.), Sustainability in Statistics Education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July 2014) Flagstaff, Arizona, USA. Voorburg: The Netherlands: International Statistical Institute.

Shaughnessy, J. M. (2007), “Research on Statistics Learning and Reasoning,” in F. K. Lester (Ed.), Second Handbook of Research on Mathematics Teaching and Learning (pp. 957–1009). Charlotte, NC: IAP.

Watson, J. M. (2007), “The Role of Cognitive Conflict in Developing Students’ Understanding of Average,” Educational Studies in Mathematics, 65(1), 21–47.

Wild, C. J., and Pfannkuch, M. (1999), “Statistical Thinking in Empirical Enquiry,” International Statistical Review, 67(3), 223–248.