PRESERVICE TEACHERS' KNOWLEDGE AND USE OF TRANSNUMERATION

Michael Christopher Daiga

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the School of Education,
Indiana University
June 2018

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

Peter Kloosterman, Ph.D.

Enrique Galindo, Ph.D.

Amy Hackenberg, Ph.D.

Crystal Walcott, Ph.D.

Yvonne Zubovic, Ph.D.

April 6, 2018

Let this dissertation be a testimony to your goodness, love and mercy. For whatever you do, in word or deed, do everything in the name of the Lord Jesus, giving thanks for God the Father through Him.

I would like to thank my wife Daren for her continued support and love through the writing of this dissertation. Look at what we have accomplished, look how far we have come. To my children Isaac, Levi and Anna, let this dissertation show you that anything is possible with the right effort and drive. Finally, this dissertation is dedicated to dad, George Juris Daiga, my first mathematics teacher.

Acknowledgements

I would like to thank Dr. Peter Kloosterman for his guidance and efforts, over the past seven years. My pathway as a graduate student was not the easiest with returning to teach in a high school classroom and my expanding family. You unwaveringly supported me throughout the process of completing this dissertation by providing constructive criticism and supportive comments. I never felt as though I do not have a place to reach out with questions, even after you retired. Many professionals view their career as a way to survive while others are a slave to a job. Few professionals are like you Pete, and use their career as a way to develop and help others daily. Thank you.

Dr. Enrique Galindo, Dr. Chris Walcott, and Dr. Yvonne Zubovic, thank you for your support and dedication to my professional development. Your impacts on my professional development have varied in role, but all contributed greatly to my present ability as a professor. Yvonne, you cultivated an early passion of statistics in my undergraduate coursework at a time in my life when I needed kindling to push my academic prowess. Enrique, you cultivated my ability to creatively train elementary preservice teachers. Chris, I am grateful for my interactions with you as a book editor and in the assessment group at Indiana University, that modeled professional stages I'm aspiring towards now. I have always felt support from you three at critical times for me personally and professionally, thank you.

I would also like to thank my AP-Statistics teachers who were expert reviewers for tasks. This research, particularly the development of tasks, was not be possible without you.

Michael Christopher Daiga

PRESERVICE TEACHERS' KNOWLEDGE AND USE OF TRANSNUMERATION

This multi-method study used a seven-task survey and paired-interviews to explore preservice teachers' Statistical Knowledge for Teaching with regards to their understanding of transnumeration, a type of statistical thinking involving graphical representations where readers translate a dataset into different forms (e.g., tables to bar graphs, stem-and-leaf plots to histograms) exposing numeracy. AP-Statistics released items were used to write 16 multiple-choice tasks designed to reveal preservice teacher knowledge of transnumeration. Based on reaction from four inservice teachers to the initial tasks, eight were revised and presented in survey form to 37 preservice secondary mathematics teachers. The survey also included questions about the preservice teachers' beliefs about their ability to complete the tasks and teach the content in the tasks to secondary school students. Thirty-two of the preservice teachers were then interviewed about their solutions to the tasks and other topics relating to teaching statistics. Survey and interview data were synthesized into constant comparison tables to provide a holistic perspective of each participant's knowledge. Preservice teachers generally lacked exposure to statistical language and in cases statistical literacy, which resulted in limited descriptions of the statistics within the tasks, limited ability to describe the statistics in a graphical representation, and significant difficulty using transnumeration to interpret a specific context. Preservice teachers were more comfortable with the graphical representation portion of a task than interpreting the written portion, which often led to a false confidence or belief they understood the information being presented in a graphical representation. Preservice teachers commonly recommended teaching the content in tasks with transnumeration, even when their content

knowledge did not seem to be adequate for fully completing a task. Finally, preservice teachers who were able to articulate types of pedagogical knowledge developed responses that included demonstrating content knowledge of transnumeration in a task.

Table of Contents

Chapter 1: Introduction

Communicating mathematical ideas between individuals, including ideas about statistics, is often difficult and thus visualizations in the form of graphical representations are often used as they provide additional information to communicate pertinent information, describe situations, or even offer further evidence for arguments. Researchers describe the general purposes of graphical representations as to communicate information (Bright & Friel, 1998) and to utilize as a tool for analyzing data (Franklin, 2007; González, Espinel, & Ainley, 2011). Although the purposes of graphical representations have simple, succinct beginnings, how a reader translates a graphical representation often is uniquely complex.

One way graphical representations become increasing complex to translate is when they require a reader to use a type of thinking called transnumeration. This term (Wild & Pfannkuch, 1999) was coined to describe the thinking that takes place when new information is produced from organizing data into different forms, exposing numeracy and in-turn facilitating understanding (Shaughnessy & Pfannkuch, 2002). Transnumerative thinking is often considered part of the process one completes to interpret a graphical representation. For example, Burrill, in her invited paper at ICOTS9 described interpreting graphical representations as the action a reader takes to "reveal stories in the data including the notion of transnumeration" (p. 2, 2014).

Transnumerative thinking can take many different forms. For example, Figure 1.1 shows a histogram utilized for the Tips Task in this study that shows 60 tips a waitress received while working at a restaurant during a week. A graph reader who is transnumerating this histogram might *break* each histogram bar into individual data points exposing the numeracy in the graphical representation. The graph reader could then describe the median data point as being one of the tips from \$2.50-\$5 because 30th and 31st data points occur within this range of values. This is merely one example of using transnumerative thinking with this histogram, different transnumerative mental processes could also occur with this same histogram. For example, a graph reader could use transnumerative thinking while comparing measures of central tendency. The Tips Task used a skewed-right histogram where the median is located to the left of the mean, or the median had a lower predicted tip-value than the mean. A graph reader may visualize the numerical spread between a mean and median to think about what a typical tip amount would be for the waitress.

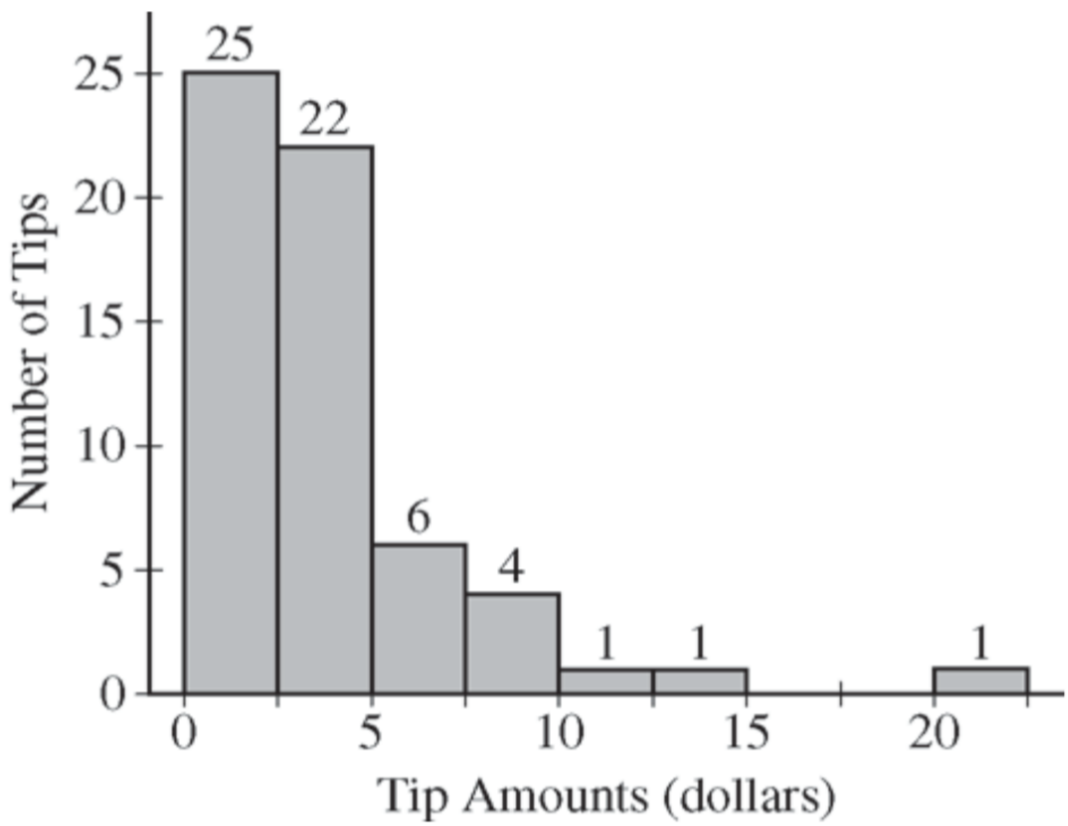


Figure 1.1 The histrogram used in the Tips Task.

Rationale for the Study

The goal of this research study is to describe how secondary mathematics preservice teachers are equipped to teach statistics using graphical representations, particularly through transnumeration. This study will describe preservice teachers Statistical Knowledge for Teaching (SKT) in terms of content knowledge and pedagogical knowledge. Because preservice mathematics teachers spend the majority of coursework on other mathematical strands, preparation in teaching statistics is often limited. In fact, the Conference Board of the Mathematical Sciences noted in their report that "of all the mathematical topics now appearing in middle grades curricula, teachers are least prepared to teach statistics and probability" (2001, p. 114). Nearly ten years later the Conference Board of Mathematical Sciences II report stated, "most new high school teachers will require further coursework to be well prepared to teach subjects such as precalculus, calculus, discrete mathematics, matrix algebra, and more than basic statistics." (2010, p. 19). Researching how our secondary mathematics preservice teachers are prepared to teach statistics, given the majority of their coursework focuses on other mathematical concepts, continues to be a national need in a complex educational system.

One resource that can help assess preservice teacher's knowledge is items used in large-scale assessments. Historically, assessments were used for different purposes in statistics education including assisting learning, measuring individual achievement, and evaluating programs (Ben-Zvi & Garfield, 2005; Pellegrino, Chudowsky, & Glaser, 2001). All three of these purposes assume that the learning process is a complete, achievable process that can be evaluated. Some assessments draw lines of pass or fail, and group participants in clusters without distinguishing the learning opportunities still present and often necessary to be successful in the future. A very different purpose of using assessments is to use individual items from the assessments as tools to assist learning. Assessment items can be analyzed for deeper, conceptual learning beyond an original item response (delMas, Garfield, & Ooms, 2005), which ironically facilitates the conceptual learning promoters of high-stakes assessments want to achieve from simply taking exams. In this research study, assessment items formed the basis of tasks that were used to determine the SKT of preservice teachers, especially the use of transnumeration with graphical representations. I specifically investigated the following research questions:

1) What do preservice mathematics teachers know about graphical representations?
2) To what extent are preservice teachers prepared to use graphical representations to help students understand statistical concepts?

Organization of the Dissertation

The following four chapters of this dissertation describe the research process from influential literature through implications of this study for future research. The literature review in Chapter 2 describes recent research that impacted large-scale assessments, statistics education, transnumeration, types of knowledge, and beliefs about statistics. Each of these topics influenced decisions during the completion of this study. Chapter 3 describes the theoretical background and methods used in this multi-method study, which required the development of tasks that were designed to influence a reader to use transnumerative thinking while answering. Chapter 4 describes results from the survey, along with specifically describing results to three tasks that were used in this research study: the Tips Task, Fuel Task, and Factory Task. Finally, Chapter 5 describes eleven explicit findings and implications of this research study and explores applications of this study across different subpopulations of our culture.

Chapter 2: Literature Review

The purpose of this chapter is to provide a synopsis of research related to the following areas (a) the development of statistical knowledge for teaching, (b) a description of graphical representation and transnumeration, (c) teachers' motivation to teach statistics, and (d) research on how assessment items can be utilized to solicit teacher knowledge. The purpose of presenting literature is twofold: to clarify the need for this research and to refine this study's research questions. The following paragraphs expand on the four purposes mentioned above.

Shulman (1986) highlighted the importance of understanding different types of teacher knowledge in various subject matters. In mathematics education, his work was a precursor to the well-known construct Mathematical Knowledge for Teaching (MKT) and the relatively new construct Statistical Knowledge for Teaching (SKT). The development of MKT in relation to Statistical Knowledge for Teaching (SKT) is described in *The Development from MKT to SKT* section. Some researchers described SKT in a broad manner, which is discussed in the *SKT Frameworks* section (Burgess, 2006; González, 2016; Groth, 2007; Noll, 2007), while others focused on specific types of knowledge or beliefs that were involved in teaching statistics discussed in the *SKT Specific Category Research* section (Batanero, Godino, & Roa, 2004; Casey, 2008; Watson, Callingham & Donne, 2008).

Graphical representations are used to communicate information. The section named *Language of Graphical Representations* defines graphical representations and the language used to describe graphical representations in the literature. Research about how preservice teachers read graphical representations, common misunderstandings of statistics found in graphical representations, and studies about transnumeration are presented in the section named *Research on Graphical Representations in Statistics.* The research on graphical representations leads into the *Transnumeration* section, where other research about the type of statistical thinking is presented. Finally, research on motivation to teach statistics is addressed in the *Teacher Beliefs on Assessments* as motivation plays a vital role in researching preservice teachers SKT.

Assessment items are an effective tool to understand preservice teacher knowledge as long as research considers the willingness of preservice teachers to answer items. For example, we see preservice teachers unmotivated to answer questions that they do not understand, which in-turn results to low motivation to teach question topics later on. Furthermore, teachers often are motivated to teach what they know or believe to be important. In part because research on statistics education is relatively new, "there has been very little research into students' and teachers' beliefs and attitudes towards statistics. (Shaughnessy, 2007, p. 1001). Therefore teacher motivation for each task was researched as it related closely to a teacher's SKT. The chapter concludes with expanded research questions.

The Development from MKT to SKT

As researchers began studying MKT, different theories of what knowledge is needed to teach emerged (Andrews, 2011; Hill & Ball, 2004; Turner & Rowland, 2011). Hill and Ball (2004) developed a theory separating mathematical knowledge into two overarching categories, subject matter knowledge and pedagogical content knowledge, which were broken into further knowledge categories. Researching knowledge of teachers in each category is a distinct focus of the mathematics education research community.

The idea that teachers need to improve content knowledge is well recognized and agreed upon by teacher groups, parents, administrators, and policy-makers alike (Conference Board of Mathematical Sciences, 2012; Ma, 1999; National Mathematics Advisory Panel, 2008).

Pedagogical content knowledge – the knowledge of how teach content – is also critical to teacher development. Often broken into subcategories of Knowledge of Content and Students, Knowledge of Content and Curriculum, and Knowledge of Content and Teaching, researchers call for careful measurement of pedagogical knowledge (Hill, Ball, & Schilling; 2008).

A purpose for studying categories of mathematical knowledge for teaching is to improve mathematics instruction. Hill, Sleep, Lewis, and Ball (2007) suggested three reasons for researching MKT: to study relationships between teachers and knowledge categories through student achievement on assessments, to study how different approaches to teacher development have different effects on pedagogical content knowledge, and to develop supporting materials for teacher education as well as professional development. Some researchers suggest that SKT is merely an extension of MKT through the mathematical strand of number and operations. (Batanero, Godino, & Roa, 2004). They discussed how professional development in number and operations often incorporates sharing a variety of possible student strategies, while professional development involving probability could also include different meanings of probability as used by students completing statistics problems. Researching statistics as a strand of MKT is merely responding to the call to develop "organization and structure of subject matter knowledge in different disciplines and what these structures suggest for teaching" (Hill, et al., 2008).

Other researchers note that the field of statistics education has fundamental differences that make it a unique, and often separated from mathematics education. The *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) report, for example, suggests an objective of statistics education should be developing students' statistical thinking through the omnipresence of variability (Franklin, 2007) which in turn includes the concept of probability – an area which is quite different from strands of mathematics that search for the exact. The role of context is altered in statistics as well, as "in mathematics, context obscures structure. In data analysis, context provides meaning." (Franklin, 2007, p. 7). Perhaps the most fundamental difference between mathematics and statistics is in how the two subjects solve problems. The statistical process of *Formulating Questions, Collecting Data, Analyzing Data, and Interpreting Results* is a methodologically different problem-solving process than other mathematics strands, therefore, some researchers (Franklin, 2015) suggest treating Statistics Education as a unique subject. In short, the differences between Statistics and Mathematics Education should spur research on the unique knowledge of needed by teachers of statistics, often called SKT.

SKT is an important research area within the field of statistics education. Shaughnessy (1992) suggested that researching SKT was one of the seven major in-need areas that statistics education needed to address, a need that was reiterated 15 years later in his chapter in the Second Handbook of Research on Mathematics Teaching and Learning (Shaughnessy, 2007). Researchers responded to his call by developing theoretical frameworks for SKT (Burgess, 2006; González, 2016; Groth, 2007; Noll, 2007) and researching specific categories of SKT (Batanero, Godino, & Roa, 2004; Casey, 2008; Watson, Callingham & Donne, 2008). However, studying SKT presents challenges because of its inexact and inferential nature (Cobb & Moore, 1997; Gal & Garfield, 1997; Rossman, Chance, & Medina, 2006; Shaughnessy 2007). These challenges influenced research method choices and resulted in a variety of different theoretical frameworks of how SKT exists, which are described next.

SKT Frameworks

Groth (2007) took research structures from the Learning Mathematics for Teaching (LMT) project (Hill, Schilling, & Ball, 2004), along with research findings on which the GAISE framework was based (Franklin, 2007) and developed a new structure describing statistical knowledge for teaching. A critical piece of Groth's framework is conceptualizing statistical knowledge for teaching into two specific *content* categories of knowledge from Hill et al.'s (2004) research: common knowledge and specialized knowledge. Groth describes interactions with common and specialized knowledge categories as necessary for many classroom activities to take place. The framework addressed how mathematical and nonmathematical categories interact, which previous research showed to be critical (Scheaffer, 2006) in achieving learning goals in statistics lessons. For example, a mean is the *mathematical* formula that is a calculation used to solve a statistics problem. But without the primarily nonmathematical activity of considering if the mean is an appropriate measure to use given the spread of data points, a statistical activity is not present. In developing this research framework, Groth (2007) recognized that his framework was hypothetical and should be added to and altered as it is compared to empirical data.

Burgess (2006) developed a framework for statistical thinking and investigating. Based on lesson recordings and interviews together with research literature ideas from teacher, mathematics, and statistics education, Burgess created a tool to investigate teacher's SKT (see Figure 2.2). The framework combines Hill et al.'s (2004) classifications of SKT in a matrix with the components of statistical thinking and empirical inquiry from Wild and Pfannkuch's (1999) research.

		Statistical knowledge for teaching			
		Content knowledge		Pedagogical content knowledge	
Dimensions of statistical thinking		Common knowledge of content (CKC)	Specialised knowledge of content (SKC)	Knowledge of content & students (KCS)	Knowledge of content & teaching (KCT)
Types of thinking	Need for data				
	Transnumeration				
	Variation				
	Reasoning with models				
	Integration of statistical and contextual				
Investigative cycle					
Interrogative cycle					
Dispositions					

Figure 2.1. Statistical Knowledge for Teaching Framework (Burgess, 2009, p. 19).

Burgess (2011) used his framework to study how beginning teachers viewed specific classroom ‘incidents’ during statistics lessons and during interviews through a variety of data sources. Burgess used four knowledge categories (see Table 2.1) to describe primary school preservice teachers knowledge in professional development to effectively teach a statistical lesson. Beginning teachers who lacked any of the four types of knowledge often missed critical statistical learning opportunities. (Burgess, 2009; 2011). Specialized knowledge of content (SKC) and common knowledge of content (CKC) were hard to distinguish, perhaps because misunderstandings of statistics were corrected as beginning teachers developed SKC.

Table 2.1.

Descriptions and Examples of Burgess' Knowledge Categories

Knowledge Category	Description of Category	An Example Using a Mean
Common Knowledge of Content (CKC)	Often considered a precursor for other categories, CKC is the content or material necessary to understand the topics being taught.	Knowing how to calculate a mean, summing data-points up and dividing by the total number of points.
Specialized Knowledge of Content (SKC)	Content knowledge teachers have that is different than other professionals.	Teachers often may think about a mean as a balance point to help explain the concept.
Knowledge of Content and Students (KCS)	Knowledge that incorporates student learning of the content.	Knowing some students struggle to visualize where a mean would be on a graphical representation.
Knowledge of Content and Teaching (KCT)	Teacher knowledge about how to teach different statistical topics, sequence topics, or plan activities that are helpful to learn content.	Understanding a mean is formally taught in 6th grade based on the Common Core State Standards.

Noll's dissertation (2007) focused on graduate teaching assistants' statistical knowledge for teaching using a task-based web survey and a series of interviews. Noll developed a framework where statistical literacy, statistical thinking, and knowledge of content and students were the three overlapping components of strong SKT. Previous research suggested that limited subject matter knowledge translated to limited mathematical knowledge for teaching (Ball, Hill & Bass, 2005; Ma, 1999). Noll (2007) found that teaching assistants had sufficient subject matter knowledge, but graduate students "did not have substantial knowledge of common student difficulties or developmental stages" (p. 319). This finding suggested the need to improve graduate students pedagogical knowledge of student development. This is a substantial finding considering that teaching assistants are new teachers, and although they likely had stronger content knowledge than elementary or secondary teachers, pedagogical deficits were at least as much of an issue as they were with the preservice teachers.

González (2016) developed a statistical framework tied closely to Ball et al.'s MKT framework, but with consideration of the subject of statistics with adjustments like renaming common content knowledge as Statistical Literacy. He argued that the elements of subject matter knowledge and pedagogical content knowledge along with teacher beliefs, and variability (Shaughnessy, 2007) are the necessary elements for a statistics teacher's knowledge. Through tasks and interviews González provided a way to assess specific component of SKT.

Haines (2015) developed a conceptualization of a framework to model Advanced Placement Statistics Teaching Knowledge (APSTK), which broke knowledge into content focused and pedagogical focused areas (see Figure 2.2). Content areas focused on exploring data, sampling and experimentation, anticipating patterns, and using statistical inference. Pedagogical knowledge areas focused on curriculum, instruction and assessment strategies, errors and misconceptions, and student thinking and learning. Haines concluded that the knowledge set of AP-Statistics Teachers was a unique blend between content and pedagogy that went beyond the typical expected knowledge of a teacher. This suggested teachers acquire a unique set of knowledge in teaching AP-Statistics that should be researched to identify knowledge deficiencies, improve preservice teacher instruction and professional development opportunities.

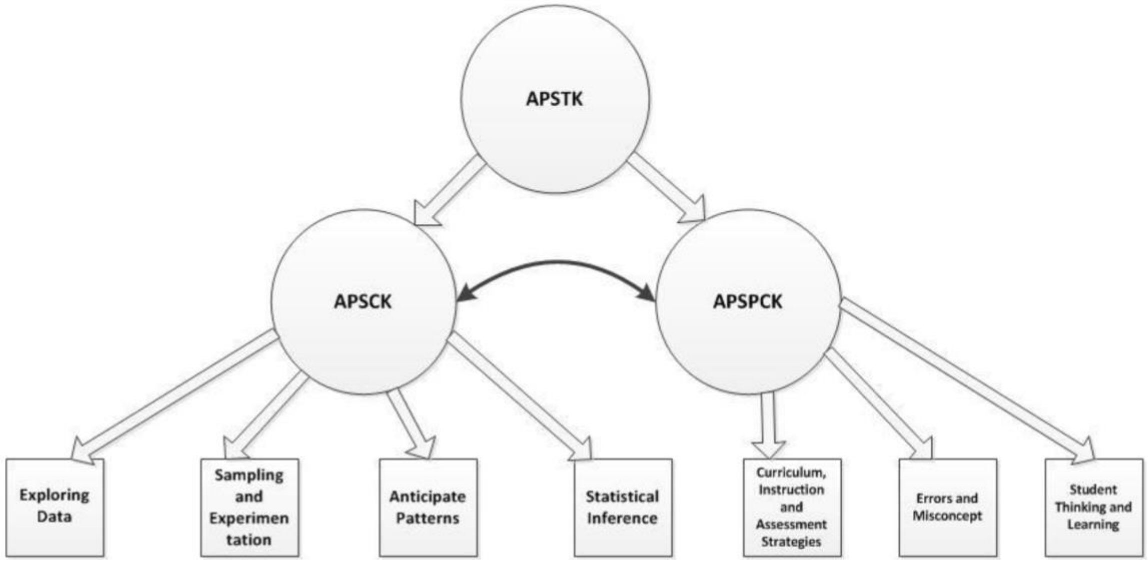


Figure 2.2. A graphical representation of the framework for AP-Statistics Teacher Knowledge (Haines, 2015, p. 10).

SKT Specific Category Research

Batanero, Godino, and Roa (2004) studied pedagogical content knowledge by analyzing primary and secondary preservice teachers' epistemology, cognition, teaching resources and techniques, affect, and interaction in two projects during a methods course. The preservice teachers were asked to use and apply the *Guide to Analyze and Evaluate the Didactical Suitability* tool (Batanero, Godino, & Roa, 2004), which was designed to improve teachers' pedagogical content knowledge through a formative cycle of reflection. Preservice teacher's struggled to utilize the tool suggesting that more instructional time is needed with respect to statistics education.

Godino, Ortiz, Roa, and Wilhelmi (2011) utilized Burgess's framework (2006) in their research on statistical pedagogical knowledge because PCK and MKT categories were "still general and could be made more precise" (p. 7). The research team had preservice teachers' reflect on descriptions of their own PCK using a formative cycle. Results suggested that preservice teachers had poor common content knowledge and poor specialized knowledge in the area of probability (Contreras, Batanero, Diaz & Ferandes, 2011).

Casey (2008) studied subject matter knowledge for teaching statistical association with secondary teachers based on the *GAISE* report. In a qualitative study, Casey classified 116 teaching incidents concerning the content knowledge of statistical association that were observed in three AP-Statistics classroom. These records of practice were then classified into twenty-nine categories that aligned with educational initiatives such as displaying solutions in multiple ways (Boaler, 2008) or utilizing technology in the classroom (Common Core State Standards Initiatives, 2010). Casey found that teachers need to have a deep understanding of both the content knowledge and pedagogical knowledge necessary to teach statistical association.

Watson and colleagues used a professional development program called *StatSmart* to help teachers consider the process students go through in learning statistics and measure teachers' PCK. Project participants were classified as having low, middle, or high PCK ability based on a set of items. However, researchers found that even teachers labeled with high PCK ability may struggle to help students' statistical development, suggesting that professional development programs explicitly target students' levels of understanding. Following the initial quantitative analysis, interview studies were completed to try to document what questioning aspects account for this teacher change (Watson, Callingham & Donne, 2008; Watson, Callingham & Nathan, 2009). Findings suggested that teachers needed to have the SCK first before PCK topics can be focused on in professional development.

González (2016) researched the SKT needed to teach variability while considering the influence of teachers' beliefs. Previously researched items were written and revised to assess

SKT through Ball's six knowledge components adapted for statistics. González proposed twelve indicators that can be used as a tool to describe teacher knowledge of variability across different contexts.

Language of Graphical Representations

The purpose of this section is to provide the reader with definitions of key terms that were used in research on graphical representation. Friel, Curcio, and Bright (2001) reviewed graphical representation research across fields and provided definitions of graphs, graph characteristics, and conceptions of graphs that were used in this study. The researchers defined *graph comprehension* as the readers' abilities to derive meaning from graphs created by others or by themselves. Graph comprehension has four critical components: purposes for using graphs, task characteristics, discipline characteristics, and reader characteristics. These four critical factors are dependent on a variety of attributes that can make the difference in reading a graphical representation. Structural components of the graph are objects such as the *framework* of the graph (i.e., axes, scales, grids, reference markings), visual dimensions of the graph called *specifiers* (i.e., lines on a line graph, bars on a bar graph), *labels* (i.e., the title of the graph), and *background* of the graph (i.e., colors, grid styles). Beyond these factors, Friel et al. suggested that different types of graphical representations have additional attributes that aide in a readers' graphical comprehension. For example, stacking "x's" in a line plot provides the reader with a representative height of frequency and in a picture graph the picture is linking the quantity of an item with this height frequency.

As important as the language that describes critical graph factors is the language that describes movements or actions of mathematics in graphical representations. The idea of *translation* is often thought of as movement within graphs. However, Roth and Bowen (2001) used translation to describe how graphs relate to situational data, or different types of graphs. *Transformations* link a visual object and an action in different subjects such as geometry or topology (Roth & Bowen, 2001). Common transformations that are discussed in mathematics are translations, dilations, rotations and reflections.

Depending on how data is presented, readers often focus on different statistical stories. For example, readers of a t-table of data may come to different conclusions on what the data is saying compared to readers of the same data displayed on a coordinate plain. As previously mentioned, *transnumeration* was a term that was introduced in statistics education to describe the phenomenon of how new information or insight can be produced from organizing data into different representations (Wild & Pfannkuch, 1999).

Tables are graphical representation themselves, but are also tools used as an intermediate step to creating other graphical representations. For example, teachers often use tables as a precursor for creating a coordinate plane. Because some students notice different statistical attributes of data in t-tables compared to coordinate planes, there is an opportunity for transnumeration to take place.

Transnumeration research is becoming more prevalent because of advancements in how technology can manipulate and display data. *Dynamic linking*, is the term that describes how connections are made by users when a purposeful action in one representation causes a reactive and coordinated action or movement in another representation (Lee et al., 2014). For example, if a user calculated a mean with one data set and was surprised by results they might think an error is present in the data. Thinking there was an error, the user might reexamine the data set by organizing it by another attribute to see if they were surprised by these further results as well. Then the user may make connections between these representations of the same data, and make conjectures about the behavior of the data set. Dynamic linking is becoming more prevalent because programs like *Tinkerplots* are allowing for quick reorganizations of data sets. Teachers use *static linking* when coordinating complementary information in multiple representations without movement (Ainsworth, 1999; Lee et al., 2014). Static linking entails a significant amount of work in drawing representations and therefore was avoided in classrooms because of time constraints. *Examining subsets* emerged in Lee et al.'s work as an action where teachers created a subset of data in describing an idea.

Curcio (1989) developed a framework that described different kinds of graphical understanding allowing the research field to focus on how *readers* of graphs translate graphical representations terming three phrases to describe how readers translate graphs: *reading* the graph, reading *between* the graph and reading *beyond* the graph. To read a graph one must understand the scale and measurement units in the graph. Reading within a graph considers changes on the graph's body, and what those changes mean with respect to the units present in the graph. Reading beyond a graph addresses inferential questions with the graph. Shaughnessy, Garfield, and Greer (1996) offered an expansion of these three levels of graphical understanding, suggesting that reading *behind* the data, or referring to using context to make connections, is also a way to read a graphical representation. Curcio used his research to fuel the education of teachers and teacher educators towards improving graphical comprehension (Curcio, 2010).

Research on Graphical Representations in Statistics

Because of the difficulty in communicating knowledge, Pierce and Chick (2011) called for research to describe teacher knowledge of key graphical features that support student learning. This section presents research describing teachers' knowledge of graphical representations. Graphical representations are a critical tool that teachers use to translate information and support communication between teachers and students. Graphical representations vary widely in style; they are commonly described as graphs, charts, or in some cases pictures. Knowledge about how teachers look at one graphical representation can be used to support researching how teachers look at multiple representations, which is necessary for transnumeration research. Because research on teachers' use of transnumeration is limited, research with other subpopulations (i.e., students, scientists) on understanding of graphical representations is included in the following paragraph.

Abrams, Oppenheim, Pazel, and Wright (U.S. Patent No. *6,285,367*, 2001) discussed a process of modeling with graphical representations as posing a question, creating a model, determining mathematical products, and deriving new knowledge. Abrams et al. suggested that students do each of these steps without a teacher lecture, using exploration and interaction with graphical representation instead. This suggestion challenges preservice teachers PCK as many of their previous experiences with graphical representations are teacher or textbook driven.

Rouan (2002) studied knowledge of statistical graphs through questionnaires and interviews and found teachers struggled to verbally extract statistical information from graphs. Rouan credited this struggle to misunderstandings in statistical content and reasoning. This research suggested that there may be a gap between teacher knowledge of graphical representations and the language teachers use to describe their knowledge of graphical representations.

González and Pinto (2008) had four secondary preservice teachers classify problems from textbooks by statistical graphs based on type of graph (e.g., bar chart) and levels of statistical thinking (e.g., reading the data). Researchers found that preservice teachers had a broad knowledge of mathematics and viewed statistics as easy-to-learn and did not consider difficulties students may encounter. Additionally, preservice teachers had little or no training in teaching statistics and considered pedagogical knowledge such as understanding stages of graph comprehension (Curcio, 1987) as unnecessary. Results suggested that finding information in between different graphical representations, which often involves transnumeration might be a curriculum gap for preservice teachers.

Cooper and Shore (2010) reported that of 75 inservice teachers (grades 4-12) in a mathematics education graduate program, 21 identified a histogram as a bar graph. Thirty-eight of the inservice teachers referred to all graphs that involved bars as bar graphs, showing that inservice teachers may not have the appropriate vocabulary to describe graphical knowledge. This research suggests that intricate graphical display ideas should be an explicit focus in professional development.

In an ethnographic study, Roth and Bowen (2001) used interviews to assess how 16 experienced scientists interpreted and used three graphs from introductory ecology textbooks. The research concluded that graphical representations do not have a significant meaning without understanding the framework, specifiers, and label meanings. The researchers suggested that a goal of graphical representations should be to depict situations in contextually meaningful ways.

Begg and Edwards (1999) studied teachers' perspective of statistics through interviews, belief questions, and concept maps. They found that teachers were confident in their ability to read charts and graphs, even though very few teachers had formal training about statistics. All teachers saw graphs as valuable to communicating knowledge (*n*=31) but only six referred to them as data exploration tools.

Kaplan, Gabrosek, Curtiss, and Malone (2014) studied common misunderstandings about graphical representations using a 10-question instrument with both forced-choice questions involving demographic questions (3) and open-ended content questions (7). The researchers found that students (*n*=350) confused histograms with case-value graphs and bar charts, that students viewed flat histograms as invariable, and that students used a time variable with histograms inappropriately. These misunderstandings persisted after instruction, suggesting they are difficult aspects of graphical representations to learn and in-turn are important to research with teachers.

Transnumeration

Recall the Tips Task example provided in Chapter 1 showed transnumerative thinking takes many different forms. For example, while respondents solved the Tips Task it was not uncommon for respondents to *break* histogram bars into data points to predict where the median would be on the graphical representation. Because transnumeration is a type of thinking that involves change, the research completed on transnumeration across the field of statistics is limited. However, use of transnumeration is steadily increasing because technology software (e.g. *Tinkerplots, Fathom*) allows easy generation of graphical representations that model real-life situations in statistical systems (Lee et al. 2011), or in other words, allows for quicker use of transnumeration. Lee et al. (2014) used *Tinkerplots* to examine teacher knowledge when investigating a statistical task. This study showed that 72% of teachers considered using one type of representation to be sufficient in answering questions; the remaining 28% used two unique types of representations. Researchers suggested that the purpose of selecting a visual display may not be based on the display's characteristics, but rather the teacher's Statistical Knowledge. Only 40% of teachers were found to make links between images, 19% of whom went from one representation to another to highlight a statistical idea not present originally (in other words, they used transnumeration).

Lee's research began to describe the transnumeration teachers utilized and highlighted how teachers struggled to use multiple, different displays to make arguments. Transnumeration between different graphical representations is particularly important knowledge for teachers because they can flexibly support student learning through multiple representations. If teachers understand how to use transnumeration in different scenarios, during student struggles, teachers can show new information using another representation demonstrating a deeper, contextual understanding of data.

Agus, Penna, Peró-Cebollero, Guàrdia-Olmos, and Pessa (2014) studied how preservice teachers' attitudes in solving statistics problems are influenced by the manner in which the problems are presented. The researchers paired problems in verbal-numerical and graphical forms and asked preservice teachers to solve them. Results suggested that assessment questions could be used to evaluate knowledge of statistical reasoning.

Arteaga, Batanero, Contreras, and Cañadas (2012) studied 207 preservice teachers' abilities in reading and creating statistical graphs. As part of a report assignment, teachers made a set of graphs. Researchers classified graphs into four semiotic levels and then had teachers' classify these graphs using Curico's graph reading levels. The research group found that teachers produced graphs advanced enough to solve tasks, however they struggled to answer the tasks correctly because they failed to use "reading behind the data" thinking.

Teacher Beliefs on Assessments

The research fields of teacher beliefs, dispositions, attitudes, and motivations are important to large spans of educational practice because of how these factors impact teaching. This research study presents research that is closely aligned to understanding how preservice teachers and assessment items interact, and the impact on knowledge from this interaction.

A critical aspect when measuring teacher knowledge is to consider teacher motivation and beliefs. Philipp (2007) recognized that teacher subject knowledge relates to teacher motivation and beliefs about that subject matter. In turn, student motivation to learn the subject matter is influenced by the motivation of their teachers. The researcher suggested two aspects of supporting student motivation: designing tasks for success and providing scaffolding for students to acquire and apply the concepts, skills and abilities of the tasks. Through elements of motivation, preservice teachers were described as having a *productive disposition for teaching mathematics,* where they were coordinating different teacher knowledge and practice skills. Through studying productive dispositions, other researchers developed assessment items to assess content knowledge, but also the respondent's opinion on the importance of item material. (Jacobson, Audeniz, Creager, Daiga, & Uzan, E. 2017; Kilpatrick, Swafford, and Findell, 2001).

Researchers studied beliefs about teaching statistics by beginning with surveys (Cashin & Elmore, 2005; Roberts & Bilderback, 1980) and using results from survey studies in follow-up research (Estrada & Batanero, 2008; Lancaster, 2007; Nasser, 2004) to describe SKT. Olfos, Estrella, and Morales (2014) demonstrated the progression of using a survey as a starting point for future research when they conducted a Likert scale survey about beliefs in what statistics is and how to teach statistics. The quasi-experimental study found that a group of teachers that observed an expert teaching lesson and discussed those lessons afterward in a lesson study changed significantly ($p > .0001$, n=28) compared to before the intervention.

Watson (2001) used a profiling instrument consisting of statistics and survey questions and 90-minute interviews with a convenience sample of secondary teachers to study beliefs about teaching data and chance. Watson found that out of nine topics, teachers were most confident in teaching graphical representations and that secondary teachers (n=28) ranked confidence in teaching graphical representations significantly higher ($p > .05$) than elementary teachers (n=15) completing the same survey.

Estrada, Batanero, and Lancaster (2011) compiled research about teacher's attitudes towards statistics and found that teachers had a variety of reasons for either viewing statistics positively or negatively. Common positive comments were that statistics was easy, preservice teachers had satisfactory statistical learning experiences, statistics is non-routine topic, statistics is useful for teachers to know, and statistics is useful in many fields. Negative attitudes toward statistics were connected with a lack of previous statistical education and the perceptions that statistical reasoning is difficult, statistics has a less formal structure than other mathematics strands, statistics is not useful in careers, and there are few statistical applications in life. The main predictors of positive teacher attitudes were previous knowledge of statistics and useful, applicable examples, and mathematics teacher educators should focus on these aspects to provide preservice teachers with the opportunity to have positive statistics experiences. Estrada, et al. suggested that future research should search for factors that influence positive teacher beliefs, particularly with secondary teachers.

Expanded Research Questions

The literature presented in this chapter points to a need for more research on pedagogical knowledge for teaching statistics. Evidence suggests that students struggle with the transition to collegiate mathematics (Conference Board of the Mathematical Sciences, 2011) making it more important to research preservice teachers' Statistical Pedagogical Knowledge for Teaching to improve prior preparation. There is still work to be done in understanding what knowledge preservice teachers need in order to provide students with classroom activities that will improve retention of important statistical concepts. Research about what preservice teachers know about the specific statistical concept of transnumeration is even more important as transnumeration often involves working with multiple representations is necessary to understand and make use of technological advances. The following expanded research questions are the focus of this study:

1) What do preservice mathematics teachers know about graphical representations?

    a) What statistical subject matter knowledge do preservice teachers display with graphical representations?

    b) What statistical subject matter knowledge do preservice teachers display with transnumeration between graphical representations?

2) To what extent are preservice teachers prepared to use graphical representations to help students understand statistical concepts?

    a) How do preservice teachers suggest they will use transnumeration as an inservice teacher?

    b) What will preservice teachers suggest students know about graphical representations? Transnumeration?

    c) What pedagogical tactics will preservice teachers suggest to teach concepts in graphical representations?

    d) What pedagogical tactics do preservice teachers suggest they should use to support students on high-stakes assessments such as an AP exam or an end of course assessment?

Chapter 3: Theoretical Background and Methods

As noted in Chapter 1, this study was designed as an exploratory study, based a situative perspective to understand preservice teacher's SKT with graphical representation and transnumeration. Chapter 2 documented that there is a limited amount of research on how preservice teachers use graphical representations and transnumeration to aid student learning. Chapter 3 is divided into two main parts: the theoretical background and methods. The theoretical background is comprised of descriptions of the Statistical Knowledge for Teaching (SKT) categories, which was used throughout the developing, collecting, and analyzing data stages of this research study. The methods section describes the techniques of research used to study preservice teacher's knowledge of transnumeration.

Theoretical background

SKT categories were a simple, yet effective way to think about and classify knowledge while completing the different parts to this research study. As noted in chapter 2, The Burgess Framework has four SKT categories: Common Knowledge of the Content, Specialized Knowledge of the Content, Knowledge of Content and Students, and Knowledge of Content and Teaching. Burgess' framework overlapped these four SKT categories with different dimensions of statistical thinking, one of which was transnumeration. Researching preservice teacher's ability to use transnumeration within the four SKT knowledge categories was a major focus for this research project.

Research questions are important, but research questions can be difficult to answer without selecting the right methods. One way this research study ensured data collected would answer research questions was by aligning data sources explicitly to SKT categories. This alignment helped justify methodological choices to developing, collecting, and analyzing data for this study. For example, when creating each task one of the SKT categories were chosen as a focus for the task. Another example of aligning data sources to SKT categories happened when designing the Interview Protocol. Each interview question was linked explicitly to a SKT category. Categories were also linked to research questions, creating an easy place to develop constant comparison table trends. The four SKT categories impacted in this study are described below.

Common Knowledge of Content (CKC).

Preservice teachers who lack CKC would almost certainly struggle to explain content to a student, or be able to describe other knowledge types involving pedagogy. This resulted in CKC questions typically being asked at the beginning of the discussion about each task. CKC interview questions asked about graphical representations, statistical literacy in the topic, and transnumeration on or between graphical representations. Writing tasks that facilitated transnumerative thinking added challenges beyond merely interpreting graphical representations. Some CKC tasks were developed from critical ideas for using graphical representations effectively, including the process of transforming information into different representations.

In theory, CKC interview questions could be asked in a meaningful way to non-teachers to gather information about the content, as CKC is held across different occupations. An example of a CKC interview question was "What statistics do you need to know to answer this question?" During interviews, CKC questions were flexible enough for the interviewees to respond in their own direction and show their knowledge. Other knowledge themes were brought up during CKC questions which complicated the development of trends in constant comparison tables, however the flexibility allowed respondents to dictate the direction of responses. Flexibility during interviews was considered a high-priority as teachers have a similar flexibility to make language-choice responses in the classroom.

Specialized Knowledge of Content (SKC).

Teachers usually have content knowledge that is different than the knowledge of non-teaching adults. For example, knowing how to calculate a mean is common knowledge that many non-teaching adults know how to compute. Thinking about a mean as a balance point would be content knowledge that teachers often know to help teach a mean, that many non-teaching adults would not know. SKC tasks were written to describe situations where it was critically important to know something about transnumeration, specific to the field of teaching (for an example, see the Fuel Task in Appendix D). Similarly, the SKC interview question were asked to understand statistical knowledge present through teaching that is not present for most adults in other fields.

The SKC interview question was asked after CKC questions on the interview protocol because they often required some of the CKC assessed at the beginning of the interview. A unique feature of SKC is that veteran teachers often take pride in describing their SKC, as it is often highlights knowledge that other professionals do not realize as important. The SKC question used in this research study was, "What math or statistics does a teacher need to know to teach this concept that a working adult in another field may not know?"

Knowledge of Content and Students (KCS).

KCS tasks and interview questions focused on preservice teachers' knowledge of students and statistical content. Tasks were often written as a scenario where a student completed a statistical question that required transnumerative thinking to correctly solve (for an example, see the Tips Task in Appendix C). The purpose of the KCS interview question was to ask preservice teachers about the appropriate age levels for content within a task. However, other interview questions intended to solicit information about other knowledge categries led into conversations about KCS because preservice teachers enjoy talking about how they will interact in the future with students. The KCS interview question was "Would a high school student see anything different from these graphical representations than you see?"

Knowledge of Content and Teaching (KCT).

KCT tasks target what statistics teachers should know to teach different statistical topics or sequences of topics. Tasks that targeted preservice teachers' KCT often had a sequence of statements or actions that could happen in the classroom or an instructional progression. For example, the Factory Task (see Appendix E) was labeled a KCT task. The Factory Task described a scenario where a high school student named Peggy made a statement and asked the reader to play the role as her teacher and choose what to do in reaction to her statement. Answer choices included reactions such as praising Peggy for her statements, asking Peggy follow-up questions, and giving Peggy an additional task to help her see deficiencies in her thinking. How the Factory Task used transnumerative content, student statements, and teacher reactions is a good example of how KCT tasks often used a sequence of events.

One of the KCT interview questions focused on curriculum design ideas, specifically with regards to how to present material for student success on the AP-Statistics exam. KCT questions were complex because often their focus was on teaching *and* transnumeration in the task. An example of a KCT question is "If you had a class of 20 students, what teaching tactics would you try to teach this concept best?" For this KCT question, the focus of the question is on teaching, but the ability to think transnumeratively through the scenario within the graphical representation was still a requirement to think about how to teach the topic.

The SKT categories above provide a basis for the intentions behind the development of each task. Another result of aligning SKT categories to data sources was explicitly considering how categories aligned with research questions. Table 3.1 aligns each research question with each SKT category. Knowledge categories from the Burgess Framework are listed across the top and aligned with expanded research questions in the left-hand column. The table contains two rankings: "Primary" represents each research questions primary purpose in describing a knowledge category and "Secondary" recognizes some research questions might be answered through data sources designed to describe other knowledge categories. Recall, interviews were semi-structured which allowed for respondents to flexibly respond to interview questions displaying different types of knowledge. Allowing preservice teachers to flexibly demonstrate knowledge while responding to interview questions aligns with the flexibility teachers demonstrate daily in interactions with students. Teacher responses to students frequently involve both content and pedagogical knowledge as both are often necessary to support student learning. One other consideration when allocating primary and secondary purposes was that limited knowledge in CKC would likely limit knowledge in other pedagogically focused categories, such as KCS.

Table 3.1

Linking Knowledge Categories and Research Questions

	CKC	SKC	KCS	KCT
Question 1a	Primary	Primary		
Question 1b	Primary	Primary		
Question 2a	Secondary	Primary		
Question 2b	Secondary	Secondary	Primary	
Question 2c	Secondary	Secondary	Primary	Primary
Question 2d	Secondary	Secondary	Primary	Primary

Note: Primary and secondary describe the main purposes of how each research question aims to describe each knowledge category.

Methods

There is a limited amount of research about transnumeration and even less is known about teachers use of transnumeration in the classroom. Therefore when designing a study to answer research questions, care was put into trying to design and use only the best tasks for interviews. As the researcher, I worked through a variety of stages, especially before administering the survey in order to provide reliable and valid data to answer research questions. Table 3.2 provides a chronological overview of the methods in this study, which will then be discussed in order through the remainder of the chapter.

Table 3.2

Chronology of this Research Study

Event	Timeframe Completed
a) Developed tasks	May 2016–July 2016
b) Sent 1st eight tasks to expert reviewers	July 2016
c) Sent 2nd eight tasks to expert reviewers	July 2016
d) Revised tasks based on feedback	August 2016
e) Used data to pick survey questions (8 tasks)	August 2016
f) Administered survey to M422 (*Teaching Mathematics in the Secondary school)*	September 2016
g) Completed M422 interviews while compiling trend notes	September 2016
h) Transcribed interviews and developed M422 constant comparison tables	Sept.-Nov. 2016
i) Administered survey to M302 (*Algebra in Secondary Curriculum)*	November 2016
j) Completed M302 (Algebra) interviews while compiling trend notes	November 2016
k) Transcribed interviews and developed M302 constant comparison tables	Nov.-Jan. 2016
l) Administered survey to M302 (*Calculus in Secondary Curriculum*)	January 2017
m) Completed M302 (Calculus) interviews while compiling trend notes	January 2017
n) Transcribed interviews and developed M302 (Calculus) constant comparison tables	Jan.-March 2017
o) Classified respondents (Alpha, Beta, and Omega), Salary Task was removed for analysis, and create new constant comparison tables by groups.	Feb.-May 2017
p) Coded transcripts by knowledge category questions, add and compare trends with survey-based constant comparison tables	March-July 2017
q) Cross-coded transcripts outside of knowledge category questions and note trends in constant comparison tables.	June 2017-Sept. 2017

Participants

Volunteer preservice teachers from three mathematics content and pedagogy courses during the 2016-2017 school year participated in this research study. Secondary methods preservice teachers were chosen because current coursework focuses on both mathematical and pedagogical aspects daily, which is similar to this study's focus. The process of recruiting participants began with the researcher observing and in some cases aiding each of the classes, which helped students become comfortable with the researcher. A few weeks into the course, the research described the study process to participants and the course instructor allotted time for students to complete the survey. Participants were told that participation in the study was voluntary and they could opt out of participating at any point in time. Participants were also told they would receive a small monetary stipend for their time if they completed the entire study, or both the survey and follow-up interview. Some preservice teachers were in multiple classes where the survey was administered. These individuals could only participate in the research study one time. This led to a sample size of 37 survey respondents 32 of which agreed to also interview. The sample was not diverse; it primarily consisted of Caucasian students. 34 respondents (92%) were Caucasian, 2 were of Asian decent and 1 was African American. The sample was relatively balanced with regards to gender, with 43% being male.

The sample ranged in the year in college based on number of credits as shown in Table 3.3 below from freshman to through a third-year graduate student. Although college credits, which closely aligns to age, was related to experience the groups formed for this study (Alphas, Betas, and Omegas) were based on the exposure to statistics or mathematics education pedagogy coursework. The primary reason for grouping in this manner was the *Guidelines for Assessment and Instruction in Statistics Education* report (Franklin, 2007) suggests exposure to statistics as a better method of describing statistical ability then age or grade in school. Because group placement was part of the analysis of the data, please see Chapter 4 for a full description on how groups were formed.

Table 3.3

Collegiate Experience of Participants by Group

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
Freshmen	4	2	0
Sophomores	2	3	0
Juniors	7	5	1
Seniors	0	1	10
Graduate Students	0	0	2

The different courses used to draw participants in this study include preservice teachers at different developmental positions in the teacher education program. One course, M422 *Teaching Mathematics in the Secondary School*, is a methods course taken by preservice teachers just before student teaching and is supplemented by a field experience. Because participants enrolled in this course are close to student teaching, their viewpoints of their training to teach transnumeration was mostly developed. The other two courses are topical (Algebra and Calculus), limited in scope (1 credit each), and are both titled M302 *Mathematics throughout the Secondary Curriculum.* M302 courses are normally taken at the same time as the corresponding mathematics classes in the mathematics department, however students occasionally enroll in M302 after mathematics courses are complete because of scheduling conflicts. The M302 courses are designed to weave teacher content knowledge and pedagogy into curriculum for preservice teachers from freshmen through junior years. Participants from the M302 courses provided a critical perspective on the development of preservice teachers' knowledge of transnumeration through exposure levels to statistics in the teacher education program.

Task Development

Using large-scale assessment items to promote learning is not a new research technique, but rather a technique that researchers have used in different ways (Groth, 2014; Jacobbe, 2007; Watson 2001). One reason large-scale assessment items support research well is their alignment with item writing principles suggested by researchers (Garfield & Franklin, 2011), which help avoid assessment validity and reliability problems. Researching a newer field like statistics education presents challenges that can complicate research. By developing tasks from AP-Statistics items, this research study leveraged aspects like topics chosen and word choice of standardized items into clearer options from the start.

Tasks were developed for two purposes: assessing preservice teacher knowledge about transnumeration and preservice teacher's ability to use transnumeration. As part of the latter, assessing preservice teacher's ability to use transnumeration, the study attempted to provide techniques and resources for other teacher trainers to use in the future and cultivate a conversation about transnumeration. In the past, researchers developed tasks from items and conducted interviews to develop professional development series (Groth, 2012; Watson, Callingham, & Nathan, 2009), or used assessment items to evaluate high school student knowledge (Whitaker & Jacobee, 2017), providing support for using assessment items as a starting point for this qualitative research. Similar to other qualitative SKT research (Groth & Bergner, 2013), this research project aimed to study both the content and pedagogical aspects of knowledge but was uniquely focused on the subject of transnumeration.

Tasks were developed for use as a tool to describe the multifaceted nature of the work of teaching. Each task was comprised of three parts: the task section, a reflection section, and a belief section. An example, the Tips Task is provided in Appendix C. Preservice teachers first complete the task and then they write notes about how they are answering the task in the reflection section. These notes provide details that support discussions during task-based interviews. The last part of each task is a series of Likert scale task-belief questions. The five belief questions were used in previous mathematics education research (Jacobson et al., 2016). The topics they addressed were opportunity to learn, ability, anxiety, confidence, and importance to teaching.

Considerations during task development.

One of the interesting, yet very challenging complications in training teachers is that besides needing to know the material (i.e., transnumeration), beginning teachers must have other knowledge (especially SKT) to teach material. Task development was complicated for a variety or reasons, one being that each task's structure needed to be carefully assembled to avoid potential distractions in word-choice, length, or problem generalization. For example, ensuring that all multiple-choice answers were similar in length avoided responders from choosing the longest answer because research has shown often longer answers are the correct answer. Rewriting an item from a focus on content to a focus on content *and an SKT category* substantially increases total word-count in the task. Additionally, final versions of tasks often involve teaching scenarios that can get complicated with the smallest changes. For example, the small difference in changing a question from whole-class instruction to single-student instruction often changes the best pedagogical choice for the situation because a teacher can focus on the specific language and methods that support one student's learning. Taking original AP-Statistics items focused on obtaining information from student responses, and changing items to obtain information about teacher knowledge was often a substantial transition. However, released information about AP-Statistics items supported this transition.

Released information from AP-Statistics assessments supported task-writing for this dissertation because such information provided background on student responses that helped teachers be more effective. Student work samples of items, suggestions for teachers' instruction, and answer key and scoring information about responses were all sources of released information that helped in task development. Often tasks were developed based on a released item with a unique graphical representation along with the suggestions for teachers' instruction that involved an idea usually supported by transnumerative thinking.

There were a handful of other advantages that large-scale assessments provided in the task writing process. The data from the large number of students that completed the items allows researchers to take characteristics such as the difficulty of an item as an attribute to begin an SKT conversation with preservice teachers. Large-scale assessment items were written to assess a specific statistical topic that was agreed upon by teams of writers. Validity and reliability of items were tested by groups of experts, which supports statistical ideas within questions. For example, a group of experts would analyze and research the word-choice used in large-scale assessment items, and when it was administered there was less risk in communicating the item.

Writing tasks with specific goals and purposes in mind is difficult without clear organization and structure. Therefore this research study used a Task Development Template, which contained two major parts; released item information and task development information.

Released item information.

The released item information part of the template (see Table 3.4) organized assessment item information. By keeping released information near each task, the intended purpose and goals of each task were revisited constantly during development.

Table 3.4

Released Item Information

Information Type	Purpose
-Purpose of Item	Provided reasoning on why the item was included in the template.
-Question identification number	Listed this number for quick referencing.
-Item details	Showed all text and graph parts of the released item.
-Correct Answer	Key provided scoring details about the item.
-Year released	Item topics shifted in statistics overtime, year released provided a perspective on the item.
-How well did students perform on this item?	Mean and standard deviation provided general difficulty information about each item.
-Teacher recommendations	Provided suggestions teachers should employ in the classroom.
-What were common student errors or omissions?	Answer trends provided teacher with information to highlight common mistakes by students.

When AP-Statistics items are released, the release includes information on the purpose of each item so that was listed at the beginning of the template. Question identification numbers, item details, and then information about how the correct answer for each item was scored appear next on the template. Often AP-Statistics items have multiple parts and overall item score is based on points for each part. Thus, correct answer sections often had a lot of conditional information about how items were scored. The year an item was released section provided additional background. AP-Statistics courses are evolving, which over the course of decades results in tests changing their focus. For example, data mining is a major statistical topic now that was not part of discussions twenty years ago. The next template section discusses how well students performed on the item with comparative measures of mean and standard deviation. These measures provided an idea of how difficult the question was for students.

The final three sections of the template provided information on student answers, common student errors or omissions, and teacher feedback. The teacher feedback section was particularly helpful in development of tasks for this study because this information suggested how improper instruction could lead to misunderstanding, and also offers pedagogical adjustments teachers can implement to improve student performance. The common student errors or omissions section helped in writing tasks by providing specific ideas for writing tasks, focused on students' knowledge of the released items. One revised task, named the Factory Task (see Appendix E) was developed for this study after considering the released exam feedback from part (e), "Students did not imply that the sample means will vary from the population means." (The College Board, 2015). Released information validated the development of the Factory Task because trends across the country showed that students struggle with differences in population and sample means.

Task development information.

The second part of the Development Template focused on information that each task needed for clarity, specifically for the revision process. The task development information portion of the template (Table 3.5) begins with the task details, followed by possible answer choices and the answer key. The rationale for including a correct answer and distractor section provided reasoning behind each possible preservice teacher response. This information was used in interviews to develop conversations when preservice teacher choose incorrect responses.

Table 3.5

Task Development Information

Information Type	Purpose
- Task details	Written task including graphical representations.
- Possible Answer Choices	Five answer choices were included in this section.
- Answer Key	Listed correct answers(s).
- Rationale for correct answer and distractors	Provided details behind the reasoning of why each answer and distractor was written.
- SKT Category	Listed the intended SKT category that the task targeted for respondent thinking.
- CCSSM, MP, or GAISE links?	Provided a specific place to link the task to important educational documents.
- Additional Interview Considerations	Provided a place to note interview questions that could be asked for each particular task.

The remaining parts of the Task Development Template involve the foci of the item. SKT categories provide the primary and sometimes secondary purposes of each task. The Common Core State Standards in Mathematics (CCSSM), Mathematical Teaching Practices, and GAISE links sections linked tasks to major resources for teachers, further validating each task's purpose. Finally, the additional interview considerations section provided a place to write notes that could be helpful during task-based interviews. Overall, completing templates for each task helped not only in rewriting items into tasks, but to structure information in a manner that will simplify the data collection and analysis parts of this research.

Although over 30 potential items were considered as possible items to develop into tasks, 16 were revised well enough to be sent to expert reviewers who provided feedback for revisions and offered suggestions to best solicit preservice teacher's transnumeration SKT. The eight tasks that are considered most relevant or useful in obtaining meaningful information about transnumeration were selected for the survey of which seven tasks were included in the entire analysis process. Below is an explicit description of the expert review procedures used for developing the tasks in the survey.

Expert Review of The Survey

Four individuals served as expert reviewers during AP-Statistics task development. Tasks were sent for review to three high school Advanced Placement statistics teachers who had an average of 24 years of experience teaching mathematics and 15 years of experience specifically teaching Advanced Placement statistics. The fourth individual was an Associate Professor of Statistics. AP-Statistics' released items that had a visual representation were considered as potential items to assess four categories of teacher knowledge: Common Knowledge of Content (CKC), Specialized Knowledge of Content (SKC), Knowledge of Content and Students (KCS), and Knowledge of Content and Teaching (KCT). Of the thirty items developed into potential tasks, sixteen were distinguished as most propitious to facilitate transnumerative thinking in respondents and were sent to reviewers. Reviewers were asked to rank how closely tasks targeted transnumeration (2-Transnumeration Required, 1-Transnumeration Helpful, 0-Transnumeration Not Needed), what teacher knowledge categories each task focused on (CKC, SCK, KCS, KCT), and how the reviewer believed interviews with each task would be in determining what a student understood about transnumeration (A Great Deal-3, Quite a Bit-2, Somewhat - 1, Not at all-0). In addition, reviewers were asked which of the six research questions could be at least partially answered by using each task. Spaces for comments were provided for each task.

Tasks were distributed to expert reviewers in two phases to limit the amount of time spent on evaluation in one sitting. Evaluation of each phase of tasks took reviewers between 45 minutes and 2 hours to complete. Comments from reviewers resulted in editing the wording of the tasks for better understanding, decisions about which tasks contained the appropriate CKC for this study, shortening of tasks, and in some cases revision of the answer key based on how the reviewers taught the material to their students. Although editing tasks typically involved minor changes, a few tasks underwent major revisions in order to be utilized in the survey.

Of the sixteen propitious transnumeration tasks, eight were selected for the final survey based on the expert reviewer rankings, and a holistic view of survey. Tasks from the first phase of review that were selected for the survey were the Typhoon Task, Mr. Sheldon's Task, the Factory Task, the Salary Task and the Hiring Task. Tasks from the second phase of review that were selected for the survey were the Real-Estate Task, the Fuel Task and the Tips Task. Although the Salary Task was included in the survey, the task was removed from analysis because of poor performance and a limited ability to answer the research questions of this study *after* it was administered across all three classes. Therefore analysis calculations only include the seven remaining tasks (see Developing the Final Survey section for more details). The following section provides an overview of the questions expert reviewer were asked, along with the reviewers' answers.

Expert reviewer questions and feedback.

The first question each reviewer answered about every task focused on whether transnumeration was necessary to answer the task correctly (see Figure 3.1).

1. Recall transnumeration is a type of statistical thinking. How much does someone completing this task have to use transnumeration?

__ Transnumeration is required to answer this task.
__ Transnumeration is helpful, but not required to answer this task.
__ Transnumeration is not needed to answer this task.

Briefly explain (if necessary):

Figure 3.1. Question about the need for transnumeration for each task.

Of the 64 responses across all 16 tasks, rankings resulted in the majority (n = 41, 64%) of tasks considered that transnumeration *was required* for the task, with the remaining amount mostly being considered transnumeration *is helpful* for the task (n = 19, 29.7%). These results were not surprising given the researcher began with released AP-Statistics items that appeared to be helpful in researching transnumeration. Table 3.6 shows the compiled results of expert reviewers, sorted by the highest mean ranking.

Table 3.6

Task Transnumeration Ratings by Expert Reviewers

Task	Expert 1	Expert 2	Expert 3	Expert 4	Mean
Typhoon[a]	2	2	2	2	2.0
Real-estate[a]	2	2	2	2	2.0
Mr. Sheldon[a]	2	2	2	2	2.0
Fuel[a]	2	2	2	2	2.0
Factory[a]	2	2	2	2	2.0
Hiring[a]	2	1	2	2	1.75
Video Game	2	2	1	2	1.75
Blue Jay	2	1	2	2	1.75
Salary[b]	2	1	1	2	1.5
Tips[a]	2	1	1	2	1.5
Job	2	0	2	2	1.5
Gasoline Tax	2	1	1	2	1.5
Stan's Support	1	2	2	1	1.5
Tree Diagram	1	1	1	1	1.0
Stem & Leaf	1	0	1	1	0.75
Fish Tank	2	0	1	0	0.75

Note. "2" in the data columns indicates required, "1" indicates helpful, and "0" indicates not needed. [a] notates selected tasks for final analysis. [b] means the task was administered in the survey, but not included in all analysis stages.

The second question each reviewer answered addressed what knowledge categories the task addressed the most (see Figure 3.2). Reviewers were asked "What SKT category does this task address the most? If you believe it addresses multiple categories, check multiple boxes AND explain why you checked each category." Knowledge categories were provided for reviewers at the beginning of the review for their reference. An open-response area was again provided for explanations.

2. What SKT category does this task address the most? If you believe it addresses multiple categories, check multiple boxes AND explain why you checked each category.

___ Common Knowledge of Content ___ Specialized Content Knowledge

___ Knowledge of Content and Students ___ Knowledge of Content and Teaching

Briefly Explain (if multiple boxes checked):

Figure 3.2. Question about the type of SKT each task required.

As explained in chapter 2, the categories of SKT can overlap making differentiation between them difficult. With that in mind, the ranking by the reviewer's ratings of SKT for the tasks are shown in Table 3.7, which provided an overview of what knowledge categories were being noticed in different tasks. One way tasks were evaluated with SKT rankings was sorting tasks by the most common responses and noting when over half (or a majority) of respondents mentioned the same knowledge category. Tasks that fit this majority response criteria, particularly with KCS or KCT knowledge were considered to provide a desired direction for this survey and interview research.

Table 3.7

Task SKT Ratings by Expert Reviewers

Task	Expert 1	Expert 2	Expert 3	Expert 4	Most Common Response(s)
Typhoon[a]	CKC	CKC, SCK KCS, KCT	KCT	SCK, KCS, KCT	3-KCT
Real-estate[a]	KCT	CKC, KCS	KCT	SCK, KCT	3-KCT
Mr. Sheldon[a]	CKC	CKC KCS, KCT	KCT	CKC, KCS, KCT	3-CKC, 3-KCT
Fuel[a]	SCK	SCK	KCT	SCK	3-SCK
Factory[a]	SCK KCS	CKC KCS, KCT	SCK	SCK, KCS, KCT	3-KCS, 3-SCK
Hiring[a]	CKC KCS	CKC KCS, KCT	CKC	CKC, SCK	4-CKC
Video Game	SCK	SCK, KCS, KCT	CKC	SCK	3-SCK
Blue Jay	SCK	CKC, KCT	CKC KCT	SCK, KCT	3-KCT
Salary[b]	SCK KCT	CKC, KCS KCT	CKC	SCK, KCS, KCT	3-KCT
Tips[a]	KCT	KCS	KCT	KCS	2-KCS, 2-KCT
Job	KCS	CKC, KCT	KCT	SCK	2-KCT
Gasoline Tax	SCK	CKC, KCT	CKC	SCK, KCS	2-CKC, 2-SCK
Stan's Support	CKC	CKC, KCS KCT	KCS	CKC	3-CKC
Tree Diagram	KCT	SCK, KCS KCT	SCK KCS	KCS	3-KCS
Stem & Leaf	SCK	CKC, KCS KCT	KCT	KCT	4-KCT
Fish Tank	CKC	CKC, KCS KCT	KCS KCT	CKC, KCS	3-KCS, 3-CKC

Note. [a] notates selected tasks for final analysis. [b] means the task was surveyed, but not included in all analysis stages.

A few considerations are worthy of mention, the first is that the Tips Task (see Appendix C) was selected for the survey even though the reviewers were split on the SKT category where it belonged. Responses for the Tips Task only included KCS and KCT indicating that material depicted in this task targeted pedagogical knowledge but the precise knowledge category was open to debate. However, because all reviewers ranked the Tips Task knowledge pedagogical (either KCS or KCT), and because all indicated transnumeration was either helpful or necessary to complete the task, it was deemed appropriate for the final survey. The Salary Task was also selected but underwent substantial changes before inclusion in the survey. Ratings on the use of transnumeration and SKT type would most likely be different if the Salary Task was resubmitted for another review.

Figure 3.3 shows the next question expert reviewers were asked with each task. Given that paired interviews were intended to either allow participants to reevaluate or validate survey responses, asking reviewers' for their opinion about the potential to discuss transnumeration with each task was important.

3. To what extent do you think discussing this task with preservice teachers will provide information on their knowledge and use of transnumeration?

   __ A Great Deal __ Quite a Bit __ Somewhat __Not at All

   Briefly explain (if necessary):

Figure 3.3. Expert reviewers were asked the question above which information gathered was used to consider the success of interviews with this task.

The eight selected tasks for the survey again were near the top of ratings for responses potential to discuss transnumeration (see Table 3.8). Two tasks selected for the survey were not ranked in the top eight by expert reviewers: the Tips Task and the Salary Task. Both tasks, however, were still rated as useful by the reviewers and both underwent significant revision before being administered in the survey.

Table 3.8

Task Interviews about Transnumeration Rankings by Expert Reviewers

Task	Expert 1	Expert 2	Expert 3	Expert 4	Mean
Fuel[a]	3	3	3	3	3.0
Factory[a]	3	2	3	3	2.75
Real-estate[a]	3	2	3	3	2.75
Hiring[a]	3	1	3	3	2.5
Typhoon[a]	1	3	3	3	2.5
Blue Jay	3	1	3	2	2.25
Job	3	0	3	3	2.25
Mr. Sheldon[a]	2	2	3	2	2.25
Stan's Support	2	3	3	1	2.25
Tree Diagram	3	1	2	3	2.25
Tips[a]	3	1	1	3	2.0
Video Game	2	2	1	2	1.75
Gasoline Tax	2	1	1	2	1.5
Salary[b]	2	1	1	1	1.25
Stem & Leaf	1	0	1	1	0.75
Fish Tank	2	0	0	0	0.5

Note. [a] notates selected tasks for final analysis. [b] means the task was surveyed, but not included in all analysis stages.

Reviewer comments on question 3 (Figure 3.3) often had the most viable suggestions for revising the task, as reviewers provided insight into their own reasoning on answers. For example, one reviewer expanded on there ranking of the Factory Task (Appendix E) by stating, "The proposed correct answer seems a bit confusing here. As a teacher, how would I ask a student to draw a normal distribution? With what data? How would the student relate this back to the given data?" Another reviewer stated "There is a lot of potential in choice 'D' for discussing how to encourage thinking in the context of the problem. Students really have to get back to the original objective, and think about what would be the optimal graph for the factory." When reviewers provided insight into there own experiences with feedback, tasks were considered to be a better fit for this research study, and often resulted in part of the task changing. For example, while reviewing the Tips Task one reviewer offered an alternative answer for how to teach how a mean and median change in a histogram if one data point might change. This alternative was included as answer "E" of the final task (see Appendix C).

Another ranking question focused on the extent to which each task would lead to answers to the proposed research questions (see Figure 3.4). Reviewers could check one or more research questions regarding what research questions could be answered by having preservice teachers complete each task.

4. Which research questions do you think might be answered, at least partially, by having preservice teachers complete this task? Check all that apply.

    - ___ 1a) What statistical subject matter knowledge do preservice teachers display with graphical representations?
    - ___ 1b) What statistical subject matter knowledge do preservice teachers display with transnumeration between graphical representations?
    - ___ 2a) How do preservice teachers suggest they will use transnumeration as an inservice teacher?
    - ___ 2b) What will preservice teachers suggest students know about graphical representations? Transnumeration?
    - ___ 2c) What pedagogical tactics will preservice teachers suggest to teach concepts in graphical representations?
    - ___ 2d) What pedagogical tactics do preservice teachers suggest they should use to support students on high-stakes assessments such as an AP exam or an end of course assessment?

Figure 3.4. Questions the expert reviewers were asked concerning which research questions each task might answer.

Results for the effectiveness of tasks targeting research questions can be seen in Table 3.9. Tasks checked by each reviewer were listed individually and counted into the majority response column if the majority of reviewers checked the same research question. One caveat was one reviewer omitted the Salary Task, so a majority common response for this task was only two out of three reviewers.

Table 3.9

Task Research Question Rankings by Expert Reviewers

Task	Expert 1	Expert 2	Expert 3	Expert 4	Majority Response
Blue Jay	2b	1a, 2c	1b, 2b	1a, 1b, 2a, 2b, 2c	3-2b
Factory[a]	ALL	1a, 1b, 2b, 2c, 2d	2a	1a, 1b, 2b, 2c, 2d	3-1a, 3-1b, 3-2b, 3-2c, 3-2d
Fish Tank	1b, 2b, 2c	1a, 2b, 2c, 2d	2c	2d	3-2c
Fuel[a]	1a, 1b	1a, 2b, 2c, 2d	1a, 1b	1a, 1b	4-1a, 3-1b
Gasoline Tax	1b	1a, 2b, 2c	1a,	1a, 1b	3-1a
Hiring[a]	ALL	1a, 1b, 2b, 2c, 2d	2d	1a, 1b, 2a, 2b	3-1a, 3-1b, 3-2b, 3-2d
Job	2d	1a, 2b, 2c	2d	1a, 1b, 2a, 2b	
Mr. Sheldon[a]	1a, 1b, 2a, 2b	1a, 1b, 2b, 2c, 2d	2c	1a, 1b, 2b, 2c	4-2c, 3-1a, 3-1b, 3-2b
Real-estate[a]	2b	1a, 1b, 2b, 2c, 2d	2b, 2c	1b, 2b	4-2b
Salary[b]	ALL	Om	2d	1a	2-1a, 2-2d
Stan's Support	ALL	1a, 1b, 2a, 2b, 2c	2c	2c	4-2c
Stem & Leaf	1a	1a, 2b, 2c, 2d	1a	1a, 2c	4-1a
Tips[a]	2c	1a, 2b, 2c, 2d	2b	2b, 2c	3-2b, 3-2c
Tree Diagram	2c, 2d	2c, 2d	2c	2c	4-2c
Typhoon[a]	ALL	1a, 1b, 2b, 2c, 2d	1b	1a, 1b, 2a, 2b, 2c	4-1b, 3-1a, 3-2b, 3-2c
Video Game	1a, 1b	1a, 1b, 2a, 2b, 2c	1a	1a, 1b, 2b	4-1a, 3-1b

Note. a notates selected tasks for final analysis. b means the task was surveyed, but not included in all analysis stages.

Given that the six specific research questions were derived from more general research questions (see Chapter 2 *Expanded Research Questions* section), it is not surprising that task rankings often focused on multiple research questions. Of the six research questions, all questions except 2a were common responses in the Table 3.8 four times or more. Question 2a, “How do preservice teachers suggest they will use transnumeration as an inservice teacher?” was never a common response across research questions, and only listed twice by reviewers on four tasks (Factory, Hiring, Stan’s Support, and Typhoon). One possible explanation could be that this research question was the only one asking reviewers to think about *how* preservice teachers will think about transnumeration and not *what* they will know. The Job Task was taken out of consideration for use in the survey because it did not hold a majority of common responses across any research questions. Overall, the tasks selected often had multiple research questions as common responses providing a range of possibilities for preservice teachers to show knowledge on survey responses and interview sessions.

The last question asked after each task allowed reviewers to provide feedback on explicit changes to each task (see Figure 3.5). Common responses included specific pedagogical modifications the reviewers themselves enacted when teaching this material, overall feelings about the task (e.g. “I liked this question quite a bit”), and suggestions for revision (e.g. I think in choice ‘B’ the word symmetrical should replace normal distribution).

5. Please provide specific feedback on revisions for this task including word-choice clarifications (e.g. "I did not understand this phrase . . ."), assessment limitations (e.g. "Your correct answer was also the answer with the most words. Good test taking skills led me to the answer, not knowledge of transnumeration"), math/stats education research (e.g. "Professor Marley did something similar in his research, you should look into that as a source for this question"), or other useful considerations in task development.

Figure 3.5. Question that provided an opportunity for task feedback and revision.

Developing the Final Survey

Results for expert reviewer responses from SKT rankings (Table 3.7) transnumeration rankings (Table 3.8) and research question ranking (Table 3.9) provided insight into which tasks were most likely to bring out preservice teacher knowledge about transnumeration. When building the survey, two ideas were considered important: what are the best tasks to facilitate answering the research questions and how do these tasks assemble into the best survey that covers a spectrum of statistics material and transnumeration? The following paragraphs provide a description of why some tasks were chosen for this research while others were excluded.

Based on comments by the reviewers and colleagues, it appeared likely that preservice teachers could complete a survey of 6-8 tasks in around 30 minutes. Five tasks (Fuel, Mr. Sheldon's, Factory, Hiring, and Typhoon) were clearly rated higher based on the three ranking categories (Table 3.10). They all had strong transnumeration scores, at least three of the four reviewers said they addressed multiple research questions, and they were seen as answerable by at least three reviewers. After these five tasks were included, tasks that had average rankings of 2.0 or higher on transnumeration were considered. The Real-Estate Task was included because of its high transnumeration score and the topic it covered was cumulative relative frequency graphs, a very different graphical representation than used in the other tasks. The Tips Task was the seventh task included because reviewers provided information on how to revise it and it was the only task that focused on how to teach the difference between a mean and median by looking at a data set through a histogram.

Table 3.10

A Combined Table of Overall Rankings

Task	Research Questions	Trans.	Majority Response(s)
Fuel[a]	4-1a, 3-1b	3.0	3-SCK
Real-estate[a]	4-2b	2.75	3-KCT
Factory[a]	3-1b, 3-2c, 3-1a, 3-2b, 3-2d	2.75	3-KCS, 3-SCK
Hiring[a]	3-1a, 3-1b, 3-2b, 3-2d	2.5	4-CKC
Typhoon[a]	4-1b, 3-1a, 3-2b, 3-2c	2.5	3-KCT
Blue Jay	3-2b	2.25	3-KCT
Tree Diagram	4-2c	2.25	3-KCS
Mr. Sheldon[a]	4-2c, 3-1a, 3-1b, 3-2b	2.25	3-CKC, 3-KCT
Stan's Support	4-2c	2.25	3-CKC
Job	No Majority	2.25	2-KCT
Tips[a]	3-2b, 3-2c	2.0	2-KCS, 2-KCT
Video Game	4-1a, 3-1b	1.75	3-SCK
Gasoline Tax	3-1a	1.5	2-CKC, 2-SCK
Salary[b]	2-1a, 2-2d	1.25	3-KCT
Stem & Leaf	4-1a, 3-2c	0.75	4-KCT
Fish Tank	3-2c	0.5	3-KCS, 3-CKC

Note. [a] notates selected tasks for final analysis. [b] means the task was surveyed, but not included in all analysis stages.

The eighth and final task that was placed into the survey was the Salary Task. Like the Tips Task, it underwent substantial revision based on feedback from reviewers. With its use of three graphical representations and an answer that was possible to uncover without advanced statistical ideas, the task was likely to promote discussion in circumstances where preservice teachers struggled on the survey. In addition, the task was released in 2016, providing a close look at a recent transnumeration topic in an AP-Statistics item. However, because no respondents actually answered the task correctly and were able to justify their answer, the Salary Task was not included in the final analysis for this study. There are several possible reasons for difficulty in answering this task. First, the Salary Task was the last task included on the survey and respondents may have been fatigued when reaching this task. The Salary Task in particular required respondents to dig into the dataset and analyze what was happening beyond the graphical representations given. Another reason the Salary Task was not included was it required respondents to look past some advanced statistical measures given on a regression output that often overwhelmed respondents. The Salary Task was also the longest task included, with its answer choices partially on the second page of the task. Although respondents did not successfully complete the Salary Task, revisions and additional research with the task may show how respondents investigate statistical questions.

Administering the Survey

The survey was conducted with preservice teachers to provide a baseline of data describing preservice teacher knowledge. There were three main sections to each survey: background information, the eight tasks, and interview logistics questions.

The first section of the survey, which asked for background information (Appendix B), included a series of questions about preservice teachers and their exposure to mathematics, statistics, and pedagogical experiences. The background information also contained a series of seven questions targeting teachers' beliefs in how successful they will be at different aspects of teaching mathematics and statistics, such as classroom management, financial stability, and ability in teaching statistics. Responses to the background questions provided insight on the differences between preservice teachers experiences, both educationally and in beliefs about future teaching abilities.

The second section of the survey included the eight tasks. For each task, preservice teachers answered the task, provided justification for their answers in a reflection, and responded to the belief questions about the task. Reflections were really important as they provided respondents the opportunity to justify answers. Additionally, respondents used the reflection section during interviews to recall the reasoning behind their answer. The series of five belief questions asked respondents to rate five statements about their beliefs on a seven point Likert scale (1-strongly disagree, 2-disagree, 3-somewhat disagree, 4-neutral, 5-somewhat agree, 6-agree, 7-strongly agree). Table 3.11 shows the five belief questions asked after each task.

Table 3.11

Third Survey Section with Belief Questions

Belief Question
1) This question is about a topic I have studied in a college class
2) I am good at answering questions like this one.
3) I often feel nervous when I try to answer questions like this one.
4) If I try hard, I can usually figure out questions like this one.
5) Secondary mathematics teachers should know how to answer this question

The final section's purpose was to organize the next stages of research. This section provided a brief description interview expectations and requested respondent's contact information. Time slots for possible interview times in the next few weeks were then listed to get an early idea of the availability of respondents. Many respondents were involved in other activities besides coursework, such as working to help pay for school. Finally, respondents were also asked what other preservice teachers they worked well with, which when combined with the availability of respondents was used to pair respondents for interviews.

Task-based Interviews

This research used a purposeful sample in part to limit stress and insecurity feelings from respondents. Pairing respondents provided a sense of camaraderie and supported discussions particularly when statistical knowledge was underdeveloped and increased the likelihood of a meaningful interview dialogue. Although respondents were occasionally paired based on interview availability because of time constraints, the sense of camaraderie was still achievable as pairings always occurred between participants within the same class.

Interviews were flexibly structured to gather data for a wide-range of teacher's reflection styles (Watson, 2001), but had a repetitive nature with the questions asked for each task. The first interview question for each task had preservice teachers discuss their answers to the task in detail. The next six questions interview protocol questions facilitated a discussion that gathered details about the four knowledge categories. The final question asked for each task provided details on why respondents chose answers the belief questions.

Interviews were typically 90 minutes long, although some interviews lasted up to 150 minutes. Part of the reason the interviews were so long was having two participants in interviews at a time required discussions of twice as many answers and perspectives as an individualized interview. Copies of each preservice teacher's survey responses were printed and provided to each respondent to help facilitate interviews. Participants were given time to review their response to a task and re-read their own reflection describing why they answered in the manner they did. Paired-interviews provided respondents with the opportunity to explain their own survey answers, but also discuss similarities and differences between their answer and the other interviewee's answer.

Task interview protocol.

The eight questions on the task interview protocol for this dissertation (Table 3.12) were chosen to help answer the research questions. The interview protocol is a flexible outline to support interview discussions, designed to help focus conversations on preservice teachers' SKT. Each written task is unique, which makes this study both interesting and challenging. The unique nature of tasks made it difficult in places to explain the trends of what preservice teachers know, and how we can improve instruction to support preservice teacher development. The interview protocol was designed to ask essential questions, but also provide flexible places for follow-up questions to fill in any gaps to describe what preservice teachers know about transnumeration.

Table 3.12

Task Interview Protocol

Interview Question	SKT category
1) Can you tell me how you got to answer?	Participant Driven
2) What math or statistics do you need to know to answer this question?	CKC
3) What information do you know from this graphical representation?	CKC
4) What math or statistics does a teacher need to know to teach this concept that a working adult in another field may not know?	SKC
5) Would a high school student see anything different from these graphical representations than you see?	KCS
6) If you had a class of 20 students, what teaching tactics would you try to teach this concept best?	KCT
7) How would you teach this concept to a high school student so they could succeed on a similar concept on an AP Stats exam?	KCT
8) Why did you respond to the belief questions about this task in the manner you did?	Participant Driven

Note. Displays link between each interview question and the knowledge category it is designed to facilitate discussion on.

Interview question one, "can you tell me how you got your answer," provided respondents with the opportunity to talk about their reasoning in arriving at an answer. Starting with a question where participants could use any of the four knowledge categories was intentional, because it allowed for respondents to dictate the importance of a knowledge category. In order to focus paired-interviews efficiently, the next six interview questions were each aligned with SKT categories and therefore research questions for this study. Interview methods were chosen because qualitative methods lend well to describing the knowledge preservice teachers have in detail. The rest of this chapter provides information about how constant comparison tables were used in this study.

Developing Tables to Compare Clusters of Comments

Preservice teacher's knowledge and use of transnumeration were analyzed with a holistic perspective, using tables in a similar manner as constant comparison tables (Corbin & Strauss, 2014). Transnumeration research across SKT is sparse, and research focused only on transnumeration was almost non-existent until this research study (see Chapter 2, section Transnumeration). Therefore a universal description of transnumeration across the literature is missing, which created challenges for researching transnumeration in a well-defined manner. An iterative process of gathering, analyzing, and developing a theory of how preservice teachers use transnumeration was closely tied to the development of comparison tables.

Developing class-based comparison tables.

The comparison tables developed in this research study developed from both survey and interview results in a woven manner, which fueled analyzing data as it was gathered over time (see Figure 3.6). In theory, utilizing a survey to gather baseline data and then using interviews to gather a deeper understanding about survey responses provided the opportunity to deeply look at how preservice teachers use transnumerative thinking within graphs. Because this research study included three classes (Methods, Algebra, and Calculus) of respondents over seven months, results were constantly being compiled into comparison tables to look for trends in the sample. Spreadsheets were created that included ten different tabs: a tab containing background information, a tab for each task (e.g., Factory tab), and a summary tab. The background tab (Figure 3.6 #1) included descriptive information (e.g., college GPA), math pedagogy and statistics course completion information, and overall Likert-scale belief question responses with reasons to teach mathematics (e.g., I believe I am teaching mathematics for a greater purpose). Task tabs (Figure 3.6 #2) included correct answers, incorrect answers, crossed out and erased answers, reflection section notes, interviewer notes, and task-specific belief question responses. After data was compiled into class-based comparison tables, initial descriptive statistic calculations were completed (e.g., respondent correct answer mean).

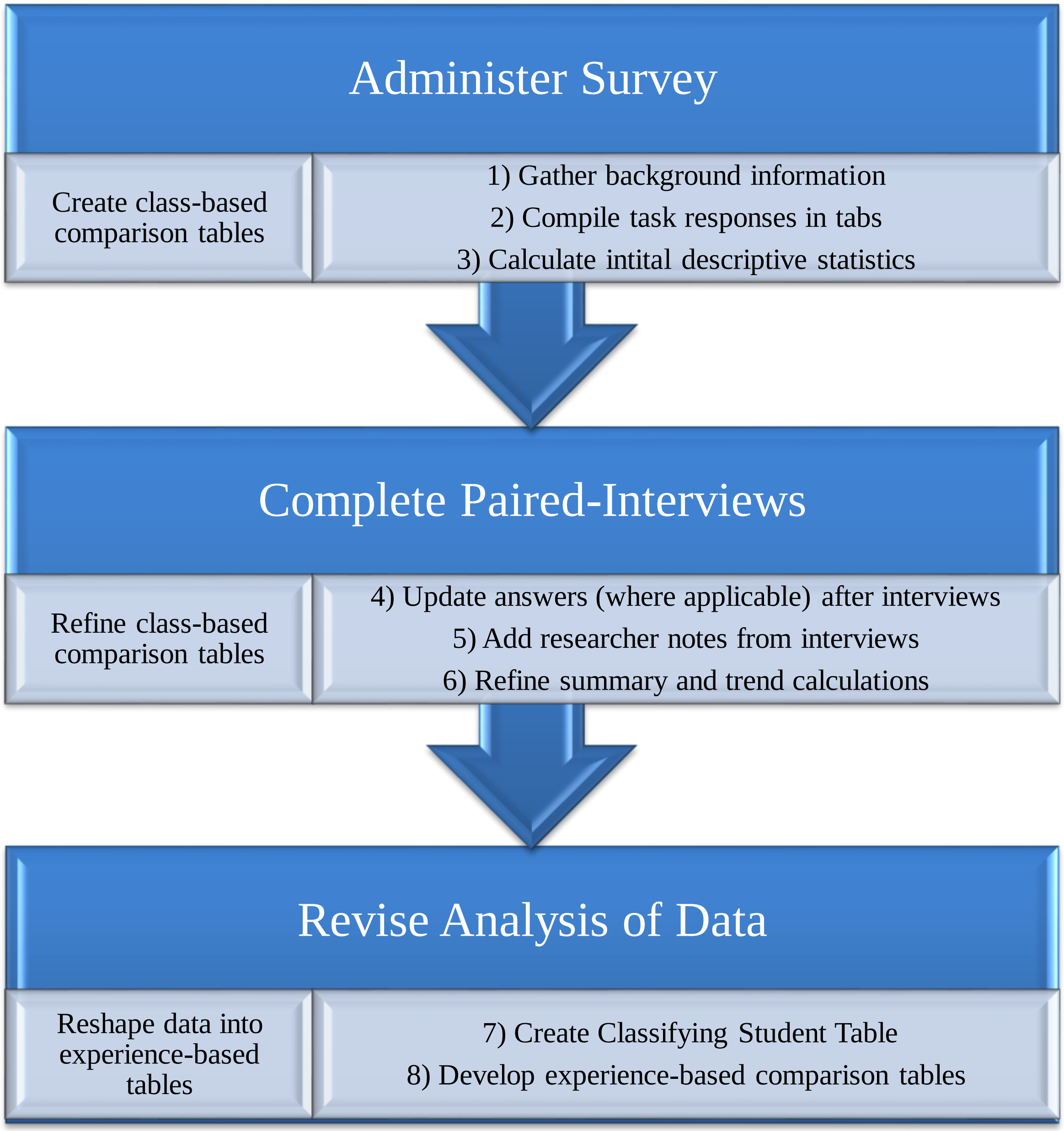


Figure 3.6. A visual of the iterative-process used to describe preservice teacher knowledge.

Recall classes where the survey and interviews were administered spanned across seven months of time. Therefore, the interviews from the first class were completed and were in the process of being transcribed when the second class took the survey. As transcriptions took place, class-based comparison tables were updated in situations where an interview respondent changed an answer when talking through the task in the interview (Figure 3.6 #4). Researcher notes from interviews were also added to provide additional details behind respondents thinking when answering tasks (Figure 3.6 #5). As interviews ended in each class, summary and trend calculations of descriptive statistics were updated in each class-based comparison table (Figure 3.6 #6).

Developing experience-based tables.

In general, respondents were theorized to have similar experiences based on the courses they were enrolled in because of course pre-requisites and co-requisites. However, the background tab of survey results revealed that both statistics and mathematics pedagogy coursework was really different within each class. For example, some students had no college statistics coursework and were taking their first mathematics pedagogy course, but had advanced notions of statistics and transnumeration because they had taken AP-Statistics in high school. Therefore a special comparison table named “Classifying Student Table” was created to compare coursework experiences across classes (Figure 3.6 #7). Evidence from the this table suggested that preservice teacher data sources be sorted based on experiences across all completed coursework, rather than merely the currently enrolled course. The choice to reclassify respondents by exposure to statistics and mathematics pedagogy over a longer period of time aligned with the GAISE framework (Franklin et al., 2007), which discussed levels of statistical ability based on experiences and not age. Reclassifying respondents resulted in a second iteration of developing comparison tables after the survey was administered to all three classes. Three groups (Alphas, Betas, and Omegas) emerged based on experience with statistics and mathematics pedagogy coursework (Figure 3.6 #8). For a description of these results, see Chapter 4 section entitled “Groups of Preservice Teachers.” These experience-based comparison tables had similar sections to class-based tables including having a background tab, a tab for all eight tasks, and a summary tab.

Selecting tasks to create three task-based tables.

The process of reclassifying data into experience groups along with respondent and interviewer reflections helped organize the selection of three tasks that told the story of how transnumeration was used across different statistical content (see Figure 3.7). The Tips Task, Fuel Task, and Factory Task were selected as tasks to further analyze and create clusters of comments (Miles & Huberman, 1994) organized by each interview question. These three tasks were chosen because interviewer reflection notes showed respondents used variety of strategies to answer these tasks, including different demonstrations of transumerative thinking. The other five tasks presented unique opportunities for respondents to show transnumerative thinking, but often had limitations (e.g., included a more obscure graphical representation).

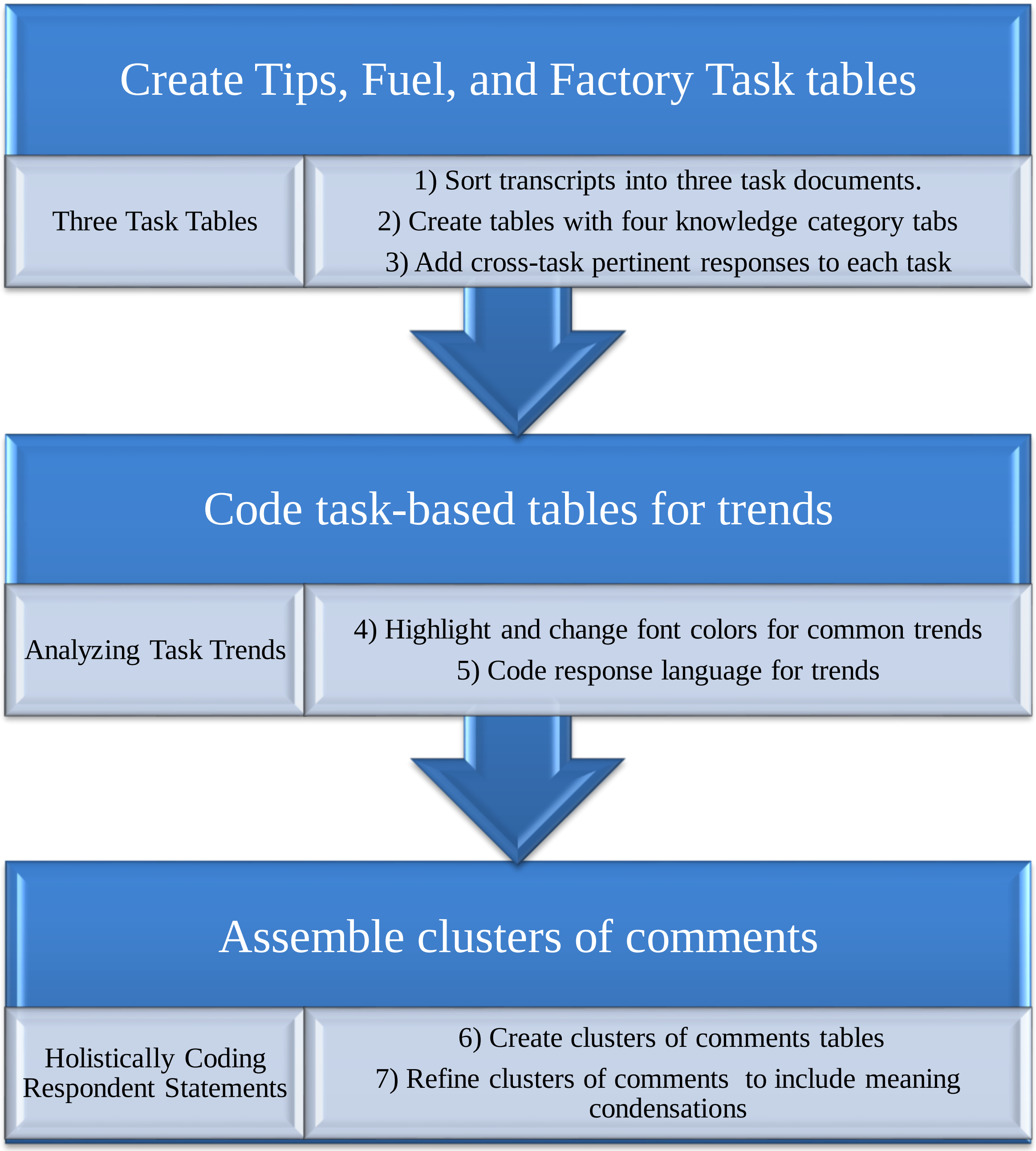


Figure 3.7. Shows process of creating task tables that led to clusters of comments in dissertation.

The process for creating task-based tables for the Tips, Fuel, and Factory Tasks began with sorting through transcripts of each interview and extracting the each tasks transcription, compiling this information into the same document (Figure 3.7 #1). These tasks documents gathered all data by task into one place, allowing the researcher to analyze the progression through answering each interview question for each specific task. Task-based tables were created for the Tips, Fuel and Factory Tasks (Figure 3.7 #2). Within each task-based table, tabs for all four knowledge categories (i.e., CKC, SKC, KCS, and KCT) were created to compile and sort interview comments. Recall, the interview protocol (Table 3.12) aligned each question to a knowledge category to help organize interview conversations for analysis. Therefore responses to interview questions were initially sorted into appropriate knowledge category tabs.

Trends the emerged while analyzing a task were added to each knowledge category tab during this process. For example, on the CKC tab for the Tips task, responses were sorted between the following emergent categories: explanation, content knowledge, graph sense, pull numbers off graph, transnumeration, mean location, median location, both comparison (mean and median), tip relationship, and change (answer) during interview based on discussion. Entire responses, but also partial responses were sorted across each of these categories as they emerged in the table. Responses were also sorted across knowledge category tabs to ensure a valid depiction of respondent knowledge. Each knowledge category tab for the Tips, Fuel, and Factory task had unique subcategories that emerged to compare across respondents, which were developed from the unique content of the task.

After this initial organization was complete for each task table, responses from the five other tasks that correspond with the Tips, Fuel, and Factory Tasks were added to each knowledge category tab to improve the holistic description of what preservice teachers know and how they use transnumeration (Figure 3.7 #3). Some tasks utilized similar statistical concepts making it important to complete across task analysis because sometime comments about one task of the interview added to the description of knowledge to another task. For example, both the Fuel

Task and the Hiring Task involved boxplots so comments that revealed respondents knowledge about boxplots during one task could provide insight on the other task.

Analyzing and coding each task table for trends (Figure 3.7 #4 & #5) was the next process in analyzing each of the three tasks. Text was both highlighted and changed in different colors to highlight correct and incorrect statements in each emergent category in a knowledge category tab. This process provided a framework to grasp common directions respondents described within their knowledge categories that could be used to discuss the major trends presented in findings. Highlighted text also provided a quick reference point for writing findings section results or providing support for statements.

The process of analyzing task trends helped organize and confirm the holistically coded respondents' statements written in the findings section. With data organized and refined in tables, sorting across tables for common trends became a feasible task. Common trends in respondent language, or groups of comments were developed into clusters of comment tables found in the findings section (Figure 3.7 #6). Furthermore, these clusters of comment tables were aligned with survey results to further develop the depiction of respondents' knowledge and use of transnumeration. The final step of this process was refining the cluster of comment tables for meaning condensations (Figure 3.7 #7). Meaning condensations (Kvale, 1996) were used to help preserve the purpose of language chosen by respondents, to better describe the situational aspects within the findings. Meaning condensations also helped assure that preservice teachers' unspoken words contributed to knowledge of their transnumeration SKT, providing for more comprehensive cluster of comment tables.

Because this research study was explorative, data was shaped and molded over the course of the research project. The development of tables is captured in Figure 3.6 and Figure 3.7 to provide the reader with a viewpoint into the lens of how transumerative thoughts were captured during the research process. Developing tables was a complex, iterative process because the project spanned three classes of preservice teachers all with unique educational experiences. The many stages of refining results through different tables supported the description of how transnumeration was used by preservice teachers. These methods provided the foundation for the following Results chapter, which describes the findings garnered from this multi-method study.

Chapter 4: Results

In this chapter, findings of the survey and interview results about preservice teachers knowledge and use of transnumeration will be presented. Based on exposure to statistics and pedagogy, similar to the GAISE report (Franklin, 2007), which categorized students into three levels, three groups of preservice teachers were created. Overall results from survey tasks are discussed along with explicit analysis of the three tasks that provided the most insight into student thinking and understanding: the Tips Task, the Fuel Task, and the Factory Task. Interview results are included throughout the chapter to provide a deeper understanding of preservice teacher's knowledge and use of transnumeration.

Groups of Preservice Teachers

University course sequences are designed for preservice teachers to obtain a degree in an orderly fashion, but often the actual course sequence for each preservice teacher varies from the intended design. Because of this inadvertent sequence, classifying preservice teachers into groups based on previous coursework was more appropriate than on their currently enrolled course (Methods, Algebra, & Calculus). Thus, the individuals that participated in the study were classified into three groups based on two factors: their exposure to statistics and their exposure to grades 6-12 mathematics pedagogy. Courses that focused on statistics or on mathematics pedagogy were considered the best coursework predictors for success on this survey because, in addition to assessing knowledge of statistics, the tasks focused on the art of teaching in the secondary school. Table 4.1 outlines the required course topics for secondary preservice teachers at the university where this study took place. Information about the number of credit hours required and information about when preservice teachers are supposed to complete courses are explained in relation to when this group of students actually took courses.

Table 4.1

Courses used to Group Preservice Teachers

	Credit hours	Designed student grade-level	Enrolled students grade-level range
AP-Statistics in High School[b]	N/A	12th Grade	N/A
How to teach Algebra[a]	1	Junior	Soph. – Grad. Student
How to teach Statistics[ab]	1	Soph,	Soph. - Senior
How to teach Calculus[a]	1	Freshman	Freshman - Senior
Middle School Math Methods[a]	3	Junior	Junior – Senior
Introduction to Probability and Statistics[b]	3	Junior	Junior – Senior
Secondary Math Methods[a]	3	Senior	Senior – Grad. Student

Note. a represents pedagogical focused-courses, b represents statistics-focused courses.

The first group of preservice teachers had the fewest experiences in statistics and teaching mathematics. The thirteen respondents in this group did not take AP-Statistics in high school and took at-most one course in college that was either focused on statistics or teaching mathematics. Nicknamed the *Alphas* because of their minimum exposure to statistics and teaching mathematics, the members of this group were relatively new to the teacher education program.

A second group, nicknamed the *Betas*, had more exposure to statistics and teaching mathematics than the alphas. *The eleven Betas* respondents either passed AP-Statistics in high school (a year-long course), or they had finished at least two courses in college involving either statistics or teaching mathematics. A college statistics course was not considered enough experience to group a preservice teacher as a *Beta* by itself because it was only one semester of material.

The final group of preservice teachers, the *Omegas,* contained the thirteen most experienced respondents in statistics and teaching mathematics. All *Omegas* were seniors or graduate students with substantial exposure to statistics and teaching mathematics. Two *Omegas* were graduate mathematics students that were considering teaching in high school in the future. All *Omegas* had taken or were enrolled in at least four courses about statistics or teaching mathematics. Table 4.2 provides a comparison of each group's years of experience in college, college GPA and high school GPA.

Table 4.2

Group GPA Results

	Alphas (n=13)		Betas (n=11)		Omegas (n=13)	
	M	SD	M	SD	M	SD
Average years in college	2.15	0.92	2.45	0.93	4.46	0.88
Average college GPA	3.18	0.62	3.24	0.46	3.34	0.44
Average high school GPA	3.91	0.48	3.82	0.33	3.84	0.33

Note. Includes mean and standard deviation comparisons of each groups experience in college, college GPA's and high school GPAs.

GPA's for each group are comparable; the most dramatic difference between groups was in number of years in college between Omegas and the other two groups of preservice teachers. Omegas averaged 4.46 years in college, in part because of the two graduate students included in the group, whereas Alphas averaged 2.15 years and Beta's 2.45 years. *Alphas, Betas* and *Omegas* had comparable group GPA means and standard deviations across both high school and college, limiting the chance that differences noted between groups throughout this research were attributed to differences in aptitude.

Overview of Survey Results

The following sections present survey results for *Alphas*, *Betas*, and *Omegas*. As explained in Chapter 3, respondents took the survey and then, on a later date, discussed their answers during the interview. Answers on the survey were often changed after they were discussed in the interview. The Salary Task, one of the eight tasks included on the survey, was removed from analysis of the data because, of the 37 total responses, only two preservice teachers answered the salary task correctly and neither of these individuals was able to provide a good rationale for what they did. Recall that expert reviewers did not rate the Salary Task highly compared to the other tasks included in the survey (see Table 3.9). The task, with major revisions, was included in hopes of learning more about how preservice teachers used transnumeration in a very different manner than the other seven tasks. The Salary Task was different because it used three different representations that included a table with raw-information about the dataset, scatterplot, and regression output. The table was theorized to provide respondents with the opportunity to explore the dataset and revise knowledge about the scatterplot, however this interaction did not occur as designed. Additionally, two answer choices about positive and negative correlation were often misinterpreted, perhaps due to the word-choice of the task. Although findings for the Salary Task are interesting, they did not help answer the research questions about graphical representations and transnumeration and therefore not included in analysis.

Overall results across the seven included tasks are presented in Table 4.3. Two different scores are presented for each group: fully correct results and *at least* partially correct results. Recall that three of the seven tasks had multiple correct answers, so partially correct meant that at least one but not all of the correct answers were indicated. Distinguishing between exactly correct and partially correct answers provides some additional insight on the extent to which preservice teachers fully understood graphical representations and transnumeration.

Table 4.3

Comparison of the Average Number of Correct Answers for Seven Tasks

	Fully Correct		Partially or Fully Correct	
	M	SD	M	SD
Alphas (n=13)	2.23	1.54	3.77	1.96
Betas (n=11)	2.18	1.08	4.36	1.57
Omegas (n=13)	2.54	1.61	5.00	1.83

Note: Group means and standard deviations are provided. ANOVA was calculated to compare group averages for both fully correct and partially correct and no difference was found. Fully correct p = 0.798, partially or fully correct p=0.236.

The average number of exactly correct responses across Alpha, Beta, and Omega groups was relatively consistent (2.23, 2.18, 2.54) with no statistically significant differences between group means as determined by one-way ANOVA ($F(2,34) = 0.227$, $p = 0.798$). Perhaps one reason for low scores was that tasks combined statistical content and pedagogical knowledge, which are both unique and difficult knowledge types to cultivate for students. (Haines, 2015; Watson, Callingham & Nathan, 2009). The average for partially correct increased across the groups with Alphas (3.77), Betas (4.36) and Omegas (5.00) although the differences between group means as determined by one-way ANOVA (F(2,34) = 1.509, p = 0.236) were not statistically significant. Standard deviations for the three groups were similar indicating there were no major differences in spread across the groups..

After each task, preservice teachers were asked to rate five statements about their beliefs on a seven point Likert scale (1-strongly disagree, 2-disagree, 3-somewhat disagree, 4-neutral, 5-

somewhat agree, 6-agree, 7-strongly agree). Table 4.4 shows the average responses from each group with regards to their beliefs about each task.

Table 4.4

Belief Results of the Seven Tasks by Group

	Alphas		Betas		Omegas	
	M	SD	M	SD	M	SD
1) This question is about a topic I have studied in a college class [a]	2.64	0.47	4.35	0.59	4.70	0.57
2) I am good at answering questions like this one.	3.62	0.59	4.43	0.70	4.49	0.57
3) I often feel nervous when I try to answer questions like this one.	4.16	0.20	4.27	0.30	4.22	0.35
4) If I try hard, I can usually figure out questions like this one.	4.32	0.39	4.78	0.45	5.24	0.53
5) Secondary mathematics teachers should know how to answer this question	5.60	0.30	5.35	0.40	6.22	0.20

Note: [a] Statistically significant difference in mean between groups. Belief statement 3 was worded so the higher the score, the more a preservice teacher felt nervous. Presumably, we would like preservice teachers to have a low score on this item.

Responses were substantially different between groups for some of the belief statements. Alphas consistently stated they studied material less in college classes, responding with an average ranking of 2.64 compared to Beta (4.35) and Omega (4.70) respondents. These differences were statistically significant between group means as determined by one-way ANOVA ($F(2,34) = 97.813$, $p < 0.001$). Given that there were seven juniors, two sophomores, and four freshmen in the Alpha group, their completed coursework should have included more of the statistics and pedagogy needed to complete tasks. Alphas believed they were not as good as the other groups at answering each task (item 2). Although not statistically significant, the increase in averages for items 2), 4), and 5) over the course of the program, there is evidence that preservice teachers believed they were developing from coursework. The fact that the Omegas had the highest average scores of the groups on items 2 and 4 indicates that confidence in doing and teaching statistics increased during the time that students were in the teacher education program. Additionally, the average score for Omegas on item 5 (6.22) indicates that these individuals felt fairly strongly that the questions posed in the survey were ones that secondary mathematics teachers should be able to answer, especially when considering the corresponding standard deviation was remarkably low at 0.20. One interesting consistency across all three groups was how nervous preservice teachers felt about tasks. All groups remained slightly above neutral (4.16, 4.27, 4.22) towards feeling nervous, showing that even teachers with substantial coursework did not feel confident in teaching statistical material in these tasks.

Overview of Interview Results

Interviews were designed to understand preservice teachers survey responses on a deeper level. The researcher paired respondents for the interviews, which included a series of questions in a semi-structured format. One interesting finding in the interviews was that, when asked to explain their answers to the survey tasks, many preservice teachers changed their responses to the tasks (Table 4.5). For example, in the first row of Table 4.5, we see that two Alphas and one Omega changed responses during interviews from a wrong response to another wrong response. The two Alpha changes took place in Mr. Sheldon's Task and the Tips Task, while the Omega change occurred in the Real-estate Task.

Table 4.5

Directional Changes of Answers During Interviews

Survey Response	Post-Interview Response	Alphas (91[a])	Betas (77[a])	Omegas (91[a])
Wrong	Alternate Wrong	Mr. Sheldon, Tips		Real-estate
Wrong	Partially Correct			Typhoon
Wrong	Correct		Mr. Sheldon[2], Factory, Tips	Factory[2], Tips
Partially Correct	Wrong		Fuel	
Partially Correct	Partially Correct	Typhoon[2]		Fuel, Real-estate, Typhoon[2]
Partially Correct	Correct	Tips	Fuel	Fuel
Correct	Wrong	Tips		
Total Number of Changes		6	6	10
Total Number of More Correct Answers		3	5	9

Note. Each line of the table displays answer changes during interviews. Interviewees changed answers both in more and less correct directions. [2] represents tasks that were changed by two participants. [a] represents the total of completed tasks in each group across all participants.

Ten changes were to completely correct answers (see wrong to correct and partially correct to correct rows in Table 4.5), whereas only two changes were from partially or completely correct answers to wrong answers. This finding suggests talking through tasks with peers can support statistical knowledge growth. Alpha respondents only changed to correct answers three out of six times (50%). Beta and Omega groups benefited from the opportunity to talk through tasks and recall previous statistical experiences. Beta respondent's changed five answers to more-correct responses during interviews, suggesting that individuals with statistics coursework were more able to figure out statistics content and pedagogy questions through discussion. Omegas followed a similar trend by changing responses ten times with the most changes towards more correct answers of any group at nine. Interestingly, only four tasks (Tips, Fuel, Mr. Sheldon, and Factory) were changed to being completely correct during interviews, perhaps suggesting that rethinking how to complete a task is an effective strategy only within certain statistical topics.

Preservice Teachers Who Completed AP-Statistics

Preservice teachers were also asked how they would teach the material in each task to help students prepare for the AP-Statistics exam. Respondents with experiences in AP-Statistics courses in high school usually had specific suggestions or used statistical language to describe what their statistics teachers emphasized while they were taking the course. There were seven preservice teachers who took AP-Statistics in high school. Trends of their own suggestions of how to prepare students for the AP-Statistics exam are shown in table 4.6.

Table 4.6

Preservice Teachers who Completed AP-Statistics Teaching Suggestions

Name	Group	AP-Statistics Exam Credit	Suggested Pedagogical Tactic
Myra	Omega	Exam Score - 2	Brought up specific suggestions for each task while using a lot of statistics vocabulary.
Jeremy	Omega	Exam Score - 2	Pay attention to vocabulary and graph. Highlighted specific suggestions based on each task.
Nick	Omega	Dual Credit[a]	Made specific content suggestions for each task highlighting what students would be doing to learn content.
Ava	Beta	Exam Score - 4	Used statistical terms through interview, content focused for each task. Asked multiple specific questions about same scenario.
Kira	Beta	Dual Credit[a]	Emphasis vocabulary and word choice, used statistical terms throughout interview. Be critical of graphical representations
Eve	Beta	Exam Score - 4	Suggested lots of practice from old exam items.
Kadi	Beta	Exam Score - 3	Use a lot of different questions that highlight the same topic in different ways.

Note. [a] Dual Credit means the student received both high school and university credit for the course.

Whether passing the AP-Statistics exam or not, there was a noticeable difference in the quality of interview responses from students who took AP-Statistics in high school compared to those who had not taken AP-Statistics. The main difference was ability to describe situations with specific statistical terms like center, spread, outliers, and so forth. Students who had not taken AP-Statistics in high school often read a task and could not interpret certain vocabulary words in the task. The Fuel Task (Appendix D) required respondents to understand what skewness was and how to apply those concepts into a teaching situation. For example, during the interview Cole, an Alpha student, said the following about skewness:

> Cole – Well, I understood skewness, but maybe not as well as the question wanted me to understand it.
> I – Ok, so what do you understand it to be? (pause)
> Cole – I am trying to put it in words, give me a moment (pause). I just can't really put it into words. But just stuff not lining-up as it should. I don't know if that makes any sense.
> I – Uh, stuff not lining-up, is it a visual thing?
> Cole – Yes maybe if you close your eyes, picture it, stuff's just (pause), I don't know. You can kind of see here (pointing to histogram) where the end of the graph is heavy over here.

Cole's struggle to articulate skewness using terminology was typical of the vocabulary struggles of most students who had not taken AP-Statistics. Myra, on the other hand was an Omega who had taken AP-Statistics in high school. She was familiar with the statistical terminology necessary to describe more attributes of skewness, even though she could not recall any specific formulas about how to calculate skewness:

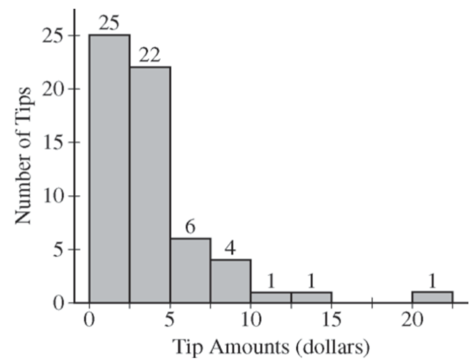
> Ok, um this one (task) I was not really sure about either because I did not remember calculations at all about skew, but the general idea of skew is that, (a distribution) leans more towards one side or the other which can cause variations between mean and median. Generally in a normal distribution, mean and median should be pretty similar, so sample mean over sample median should have a ratio close to 1. If there was skew, that would be further from 1.

Myra used more advanced statistical terminology like variation, mean, median, normal distributions, sample mean and sample median throughout her interview. One overarching theme was that preservice teachers were often confident in claiming understanding of a statistical topic (e.g., a histogram), even though they did not have the statistical background to answer the task correctly. A good example of this was Cole's confidence in understanding skewness, but not being able to describe skewness like Myra with specific vocabulary.

Tips Task results

The Tips Task (Figure 4.1) was selected to provide information on how students use transnumeration within a histogram by breaking apart or grouping data points in histogram bars and using those data points to think about mean and median calculations. The task was designed to use statistical content knowledge that secondary mathematics preservice teachers should be familiar with. One concern when proposing this research study was with their limited training, preservice teachers generally do not possess much statistical knowledge. With transnumeration being a type of statistical thinking, it was theorized that if the content itself was too challenging respondents would not be able to start the task because of the topic rather than because they cannot use transnumeration. In this scenario, very little information would be gained about preservice teachers knowledge and use of transnumeration. The main content of the Tips Task was mean and median, topics that are taught beginning in 6th grade (CCSSM; 2010). Preservice teachers had a concept of what mean and median were, but did not have the skills necessary to answer this task because of their limited ability to transnumerate the data.

Johnny is a student of yours who is working on the following problem: “The histogram below shows a waitress’s 60 tip amounts from this past week. One of the tip amounts was mistakenly added to the graph as being $8, when in fact the tip was actually $18. What effect would this mistake have on the mean and median?” Johnny responds to the question by saying, “The mean and median tip amounts should both increase.” Is Johnny correct? What would your next action be in response to Johnny’s answer?



A. Johnny is correct. Next, I would ask Johnny to describe the appearance of the distribution because it’s important to understanding the data set.
B. Johnny is correct. Next, I would ask Johnny if he thought it would be helpful to track tip amounts for another week to see similarities and differences.
C. Johnny is incorrect. Next, I would have Johnny sum the differences between each tip and the mean and compare them to each other.
D. Johnny is incorrect. Next, I would ask Johnny what effect removing the $20+ tip amount would have on the mean.
E. Johnny is incorrect. Next, I would have Johnny cut each histogram bar out and tape the bars together in order by length. Then have Johnny fold the entire length in half to show the median and discuss how this would change given the clerical mistake.

Figure 4.1. The Tips Task and answer choices, answer “E” is the correct response.

The Tips Task was a unique task for this survey in that answer choices funneled preservice teachers in two directions: when Johnny is correct (i.e., “A” and “B”) and when Johnny is incorrect (i.e., “C,” “D,” and “E”). Preservice teachers were unsure of whether Johnny was correct or not, with 24% (n=9) responding that Johnny was correct. Answer choices “C” and “D” provided critical distinctions from the correct answer choice “E.”

Table 4.7

Number and Percentage of Preservice Teachers Answer Selection by Group on the Tips Task

Answer	Alphas (n=13)	Betas (n=11)	Omegas (n=13)	Total	Percentage
A	3	3	2	8	21.6%
B	0	0	0	0	0%
C	2	3	1	6	16.2%
D	1	2	2	5	13.5%
E[a]	5	3	7	15	40.5%
A & B	1	0	0	1	2.7%
D & E	0	0	1	1	2.7%
Omitted	1	0	0	1	2.7%

Note. [a] denotes the best (i.e., completely correct) answer.

Answer choice "C" was created to use important, and frequently discussed terminology in statistics courses. Answer "C" hints towards using standard deviation with summing the differences between tips and the mean and comparing them. This distractor was intended to target respondents that did not have a deep understanding of what is involved when describing the spread of a distribution by standard deviation. Finding the typical standard deviation was actually not possible if respondents were thinking about the data point values within each histogram because only the possible range for each data point was provided (e.g., $0-$2.50) rather than specific values. Respondents who chose this answer were likely not able to transnumerate data into a form to notice it was not possible to calculate a standard deviation without exact tip values, but believed this response sounded the most statistical.

Answer "D" suggested the teacher remove another tip from the histogram that would affect the mean, but not change the median at all. The intent of this answer choice was to provide respondents with a similar action to initial question, but not affect the student misunderstanding, which focused in on how the median is changing. Preservice teachers who choose answer "D" often saw similarities to the initial change in the task and wanted Johnny to visit the same idea again to correct his mistake. Reasoning for choosing "D" often focused on the idea that students need more practice to understand this concept.

Answer choice "E", which was an added at the suggestion of a reviewer, distinguished between the mean and median. Preservice teachers who choose "E" had to think about the content knowledge necessary to understand mean and median and how that was affected differently by changes in a data set while also thinking about how to teach content to students.

Answer choices "A" (21.6%) was the most common incorrect choice when respondents believed Johnny was correct that the mean and median would both change. Answer "A" was created to provide an answer that focused on the spread of a distribution, an important statistical concept for teachers to understand. However, when considering that Johnny's statement was focused on how measures of central tendency work, discussing the distribution would not be the most helpful tactic to work at the root of Johnny's misunderstanding.

Tips Task interview discussions that led to answer changes.

The paired interviews provided the opportunity for respondents to expand on their justifications from the survey and also to improve their own knowledge. An example of development was found with Ava (a Beta who scored "4" on the AP-Statistics exam) and Suzy (an Alpha). On the initial survey, Ava answered "A" on the Tips Task, indicating that as "Johnny is correct so that means it could be answers "A" or "B". However, "B" would not help Johnny understand, so "A" is the correct answer." Suzy initially answered correctly with "E" and justified her choice by writing "the median does not change, because 1 higher tip will not have an effect on the 25 and 22 lower tips." In the excerpt below, Ava and Suzy were discussing how they arrived at their respective answers:

Ava – So the first thing I wanted to decide in my head was, is he correct or incorrect, because that would eliminate three answers or two answers. So I thought he was correct, because the mean would increase because you are adding in a number that is bigger than before so it would cause the data to increase, and then the median I thought would also…(pause) increase? Because the data is like (pause), well I was kind of debating on the median, because a lot of the data is over to the left-hand side. So I thought it might move it over by one to the right-hand side, which would increase it.

Suzy – I was debating on the median also. I came to the same conclusion about the mean increasing, but in the end I said the median would not increase just because it's only adding like one over here. I wrote on mine like this one would go down to three, right? And this one would go up one?

I – Ok, the histogram bar with four goes to three and then you drew a one over by the $20.

Suzy – Yes, so I think just the one number changing would not affect the median I did not think. Kind of like I said in the other question, if you were to write out all of the numbers and then cross them off from the beginning and end, I think you would end up at the same number. But I could be totally wrong because I did not actually do that.

Ava – Well, looking at your graph and the way you drew it in, that definitely makes sense. Yes, I could definitely see that. And even if it did increase, it would be really minimal.

Suzy – Yes. (pause) Yes, I don't know. I think I would actually have to write it all out which I did not do to answer this question.

I – Ok, is that kind of how you feel too Ava? If you wrote it all out, you feel like you would figure it out?

Ava – Yes, it would definitely be a lot easier than just conceptually thinking about it.

Both respondents seemed to have a relatively limited understanding of a median. Ava and Suzy were both comfortable finding a median with a set of numbers, lining numbers up from least to greatest. Ava even mentions her limited knowledge of the median, stating that lining all numbers up from least to greatest and crossing them off one at a time would "be a lot easier than just *conceptually* thinking about it." To solve the Tips Task, respondents had to conceptually group data together to find the median, which was a different task than most preservice teachers' previous experiences. Through this conversation, Ava benefited from Suzy talking about how she would line numbers up and find the median, where initially Ava was closely linking how median and mean function.

Correct solutions justifications for the Tips Task.

Of the 15 respondents that answered this question correctly, 14 respondents answered the correct response of "E" during the survey stage. Many respondents were confident that "E" was correct, but struggled to justify their answers with statistical language and accurate information about measures of central tendency. The following section provides information about preservice teachers who answered correctly answered "E" and how confident they were in completing the Tips task organized by group.

Table 4.8 shows the Alpha respondents' beliefs about whether they were good at answering questions like the Tips Task and provides the primary justification each respondent made for answering "E." Respondent justifications came from both reflection section notes and interview questions.

Table 4.8

Correct Alpha students justifications for Tips Task

Respondent	Good[a]	Justification
Klara	4	Klara chose "E" because she liked visual things and because "the waitress had a lot more tips of $5 or less then she had of 5 or over. So then seeing that I would think, my median is probably going to be in that group because … most of them are."
Opal	6	Opal said "because the median is somewhere between 2.5-5, based on the histogram, changing $8 to $18 wouldn't change the median. It would only change the mean."
Jill	5	Jill did not interview but her reflection said, "I think this next step provides the most understanding into Johnny's mistake."
Suzy	5	Suzy wrote "the median does not change, because 1 higher tip will not have an effect on the 25 and 22 lower tips."
Mia	1	Mia wrote "Johnny is incorrect because of the abundant of no tips, this wouldn't change the median, just the mean." Although Mia answered correctly, she stated incorrect information about reading the histogram.

Note: [a] Good column is the response to statement, "I am good at answering questions like this one" where 1 is strongly disagree and 7 is strongly agree.

Klara provided the desired response to the Tips Task by making statements showing she had a visual understanding of how mean and median were being affected by the change in the data, and by making pedagogical statements that considered a variety of student abilities. For example, when asked how she would teach this concept to a group of twenty students, Klara said:

> I do like pencil and paper, but one thing I have started when I am with kids is a lot of times the way you learned is not the best way for other kids to learn. That's why I try to think of better ways, I know visual things work really well for kids who don't know math. Just from experience, that's why I picked that (answer "E").

Mia was an example a respondent answering the Tips Task correctly, but not understanding how to break apart the task into meaningful content to answer the question. For example, Mia did not know what the “6” and the “4” above the histogram bars were representing, and did not show any evidence of being able to break apart the dataset to compensate for the misplaced tip amount. Mia showed little if any evidence of transnumerating the dataset, and strongly disagreed that she was good at answering questions like this one. Three of the 11 Beta preservice teachers responded correctly to the Tips Task, but had varying levels of understanding of both the content knowledge and pedagogical knowledge desired to answer the question (Table 4.9.). All three Beta respondents who responded correctly rated themselves at 5 or above indicating that they were confident that they were good at answering this question. Taft declined to interview, but provided a description in his reflection section that showed understanding of the influence on the median.

Table 4.9

Correct Beta Students Justifications for Tips Task

Respondent	Good[a]	Primary justification of the respondent
Ben	5	Ben debated between answer choice C, D and E during the interview saying “I think Johnny is incorrect. I think the mean would increase, but I do not necessarily think the median would increase.”
James	7	“No need to address A and B, since Johnny is incorrect. No need to talk about mean b/c mean does increase (coincidentally by $.17). ‘E’ targets the median, which does not increase.”
Taft	5	Taft did not interview but her reflection said, “I said Johnny was incorrect because the mean would slightly increase but the median would not move because it is already above the median, so it would not move.”

Note. *a* Good column is the response to the belief questions, “I am good at answering questions like this one” where 1-strongly disagree and 7-strongly agree.

Although Ben answered correctly, his knowledge of the Tips task was limited in places, particularly with how mean and median were calculated without explicitly supplying a dataset.

For the interview, Ben was with Cole who responded incorrectly to the Tips task and stated during the interview that he had "no clue" on how to answer this task. The following transcript provides information on the mathematics and statistics Ben noticed in the Tips task:

> I – What math of statistics do you need to know to answer this question?
>
> Ben – How to find the mean and median. (pause)
>
> I – How do you find those in a graph like this? (pause)
>
> Ben – I mean I guess you can't technically find them, but you can roughly estimate that the mean is going to be low because the tip values are low. And the median is probably going to be lower too.

Later in the interview, Ben provided more information about his knowledge of mean and median when talking about what we can see from the graphical representation:

> Ben – I kind of like "D" as an answer too because if you remove the $20, you can see that takes the mean down a lot. If you added an $18 instead of an $8 tip you would see the mean would increase a lot, so I kind of like "D" as an answer now that I think about it.
>
> Cole – Because "D" is doing the same question, and I like the question.
>
> Ben – "D" is doing what the question did, only in reverse.

From this transcript it was clear that Ben can look at a histogram and consider the change a single data-point may have on the mean value of the dataset, ruling out a limitation in understanding how the mean was affected. Ben later stated he believed working adults "have a better grasp on what mean is then median, because I feel like most people know mean is average, but then median kind of throws people off" further suggesting that his own understanding of how a median works in a distribution was limited. Despite the limitation of knowledge on the median, Ben was still confident in his answer stating at the conclusion of our discussion about the Tips Task that, “I definitely don’t think I have seen anything like this in college, but I think that’s because it’s a simple enough thing that college does not really need to teach you that.”

Omega respondents did better than the other groups on the Tips Task, suggesting greater ability than the Alphas and Betas to think pedagogically while using transnumeration. All correct Omega respondents were able to articulate differences between the mean and median with regards to the histogram as a visual or the numerical values given in the dataset. Correct Omega respondents except for Myra believed that they were good at answering questions like the Tips task (Table 4.10). Myra hesitated to answer positively because, although she answered “E”, she interpreted lining-up the histogram bars by largest total number of tips first, not in order of dollar value of least to greatest. Lining up histogram bars by the largest amount would be incorrect for some datasets, so Myra’s interpretation of answer “E” had valid reasoning behind her response. Myra did explicitly state the median would stay the same and the mean would change and suggested asking Johnny how the mean and median were computed followed by guiding questions until Johnny realized his mistake.

Table 4.10

Correct Omega Students Justifications for Tips Task

Respondent	Good [a]	Justification
Carl [b]	5	Believed both "D" and "E" could be useful, but knew that if Johnny did calculate removing a $20 tip he would not see a change in the median. Believed not seeing a change could help reveal how a median is affected by changes in a dataset. Carl was nervous in belief responses because mean and median together are tough.
Nick	7	Said "E" is the only fact, most are good questions though. Described movements of dataset in detail with regards to mean and median.
Sean	5	Stated the median would stay in the range ($2.5-5) because many data points lie in that range.
Kai	5	Mean would change, but median would not because there was so much data below 5.
Jeremy	5	Chose "E" because he is wrong the median would not change
Chloe [c]	4	Left a blank reflection and did not interview but wrote, "AVE" by mean and "MID" by median in question
Myra	3	Chose "E" but want to proceed differently. I would ask Johnny how mean and median are computed and continue to ask guiding questions until he realized his mistake.
Kent	7	Initially answered incorrectly I, changed answer during interview to "E." After interview Kent believed answer "C" would be helpful in understanding situations with a mean, "E" with a median.

Note. [a] Good column is the response to statement, "I am good at answering questions like this one" where 1 is strongly disagree and 7 is strongly agree. [b] Carl was partially correct answering D and E. [c] Chloe was a 3rd year graduate student.

Kai was paired with another Omega preservice teacher, Kent, who was partially correct on all seven-survey tasks and was often the first and most active discussant during their paired-interview. Kent answered incorrectly initially on the Tips Task responding with choice "C" and later changing his answer to "E." Kai helped Kent see how the mean and median were affected differently by one change, justifying her answer choice by saying, "answer 'E" is targeting what

Johnny got wrong about in his statement. "E" is making him see that the median would still be the same and that's what was wrong in his response so I think that's what needs to be focused on." Later on after Kent had changed his answer, Kent and Kai were asked how to best teach this concept to a class of twenty students and had the following conversation:

> Kent – I think you could do an activity almost, with physically having them map out the tips. If it was not tips, something more physical and you can show exactly why the claims are true or why Johnny is incorrect.
>
> Kai – I kind of like the suggestion (in answer "E"), obviously it probably would not be the best to use with 20 kids, but maybe if they were in (smaller) groups, having them cut-out the bars. This length is 25, this length is 22 and so when they can fold it together they can see where the median can be. Students can even chop off the first 20 values to see how the median would still stay the same…
>
> I – Why do you think you like that idea Kai?
>
> Kai – I think that sometimes it can be hard when the median is asked for in a histogram because students can be like "ok its got to be the middle of the numbers", even though this histogram has a lot of data here (pointing at 25), they may not always make that connection that this is a big portion of the distribution and so the middle is going to be somewhere in here.

Kai showed relatively advanced knowledge of how a median works particularly when suggesting to "chop off the first twenty values to see how the median would still stay the same." Many respondents could not balance data-points in the histogram comparatively on each side of the median. Kai's suggestion was even more advanced, or perhaps more confidently stated, choosing the large number of twenty points to remove and visualizing the center of the dataset must focus in on the lower tips. Knowing that there were twenty-five values from $0-$2.50 and thirteen values between $5-$22.50 out of sixty total values leaves the median value in the $2.50-$5 histogram bar. Where many respondents were not confident in their comments about whether the median would change, (e.g. "the median may change a little"), Kai understood and explained thoroughly why the median would not change in this dataset.

Common Knowledge of the Content and transnumerative thinking.

Previous research suggests that knowledge of measures of central tendency is often limited to simplistic or incomplete definitions (Jacobbe, 2011). For example, a median is often defined as the "middle number" and people struggle to calculate the median without lining numbers up from least to greatest and crossing off numbers one at a time. For the Tips Task, understanding the influence of a change to the dataset on both the mean and median were critical ideas to arriving at answer "E." There were instances where preservice teachers answered incorrectly on the survey, but interview conversations facilitated the transnumerative thinking necessary to understand why "E" was correct. Will, an Alpha, was one of those who initially answered "A" but after his interview partner explained why the median does not change he stated, "I got the mean and median part wrong." Table 4.11 specifically shows statements that respondents made on survey reflections and interviews about the mean and median and how changing an $8 to $18 tip would affect these two measures of central tendency. Respondents who changed their view on whether the median would increase or stay the same from survey to interview were noted, although this did not always result in a change in a correct answer choice nor was it the only reason respondents changed answers on the Tips task.

Table 4.11

Evidence of Common Knowledge of the Content on the Change of the Mean and Median

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
Stated "I don't know"	*Neil, Bell* c, *Cole*		
Mean & median increase			
-Mean & median increase	*Jay, Avis, Will* a,	*Sally, Eve* a	*Julie, Kent* a, *Hope*
-Median not necessarily or slightly change		Ben, *Kira*, *Amy*	*Mary,*
-Cross off numbers till median	*Lisa*	*Ava* a	*Mya* c,
Median stays the same	Jill c	*Ava* b, *Chas, Kadi, Jack,*	Chloe c, *Carl*
Transnumeration with the median			
-Grouped a cluster of data under $5	Klara, Opal, *Will* b,	James	Kai, *Kelly*
-Visualized $8 and $18 on right hand side of median.	Suzy,	Taft c	Kent b, Myra, Jeremy
-22 bar contains the median	Mia, *Dean*	*Eve* b	Nick, Sean

Note. a this was the respondent's survey response. b this was the respondent's interview response. c this respondent only took the survey. An *italicized name* indicates the student answered the Tips Task incorrectly, a regular font name indicates the student answered the Tips Task correctly.

Preservice teachers ability (or lack there of) in visualizing how a mean and median lie in a histogram either were key to answering the Tips Task. Not surprisingly, the only respondents to state "I don't know" and not arrive at a clearly-reasoned answer were Alphas. One respondent, Cole, discussed how he viewed the Tips Task when completing it during the survey:

> Cole – Oh, I remember I was so lost on this (task). I could not answer it. I did not know if he (Johnny) was right or wrong.
>
> I – Oh, as far as correct and incorrect?
>
> Cole – Yes, yes, yes, I did not know if Johnny was right or wrong. I took too long in trying to figure that. I think I picked "C" because of the explanation, not because of his answer. I said "if you sum up the differences between each and compare them to each

> other," I really even in my explanation, I even said, "I don't know this question, but comparing numbers is good." So yes, just have Johnny compare numbers I guess that's what I wanted.

Respondents who stated the median would increase overwhelming answered incorrectly, with only one out of the 15 answering correctly. These preservice teachers were often only able to *read the graph* (Friel, Curcio, & Bright; 2001) by providing evidence of understanding the scale and measurement units of the graph. Some statement examples of respondents who were reading the graph focused on the total amount of tips written above each histogram bar, stating there were 60 total tips, mentioning tip amounts were grouped in ranges of $2.50, or noticing the number of tips were scaled by fives on the dependent axis. Although these statements could be helpful to understanding the dataset, they do not incorporate the transnumeration needed to solve the Tips Task at all.

Median stays the same.

Interviews revealed that both correct and incorrect respondents wavered among multiple answers and decided on answer "E" partially because it was a hands-on activity that had students active in learning. Five of the seven respondents who stated the median value would stay the same whether the tip was recorded as $8 or $18, but could not provide a specific transnumerative statement about the dataset. The two correct responses were survey-only participants with evidence provided in the reflection sections. Had these individuals participated in the interview, they may have made specific transumerative statements, which would categorize their knowledge differently in the *median stays the same* row of Table 4.11.

Transnumeration with the median.

Some respondents who *read within the graph* analyzed the graph's body and edited the graph based on the tip change from $8 to $18. For example, eight of the Tips Task respondents drew a line shortening the $7.50-$10 histogram bar and created a histogram bar for the correct tip amount ranging from $17.50-$20 (see Figure 4.2). Two of these students were Alphas, one was a Beta, and the remaining five respondents were all Omegas. Other respondents counted the given number of tips in each range to get to 60 total tips (i.e. 25+22+6+4+1+1+1).

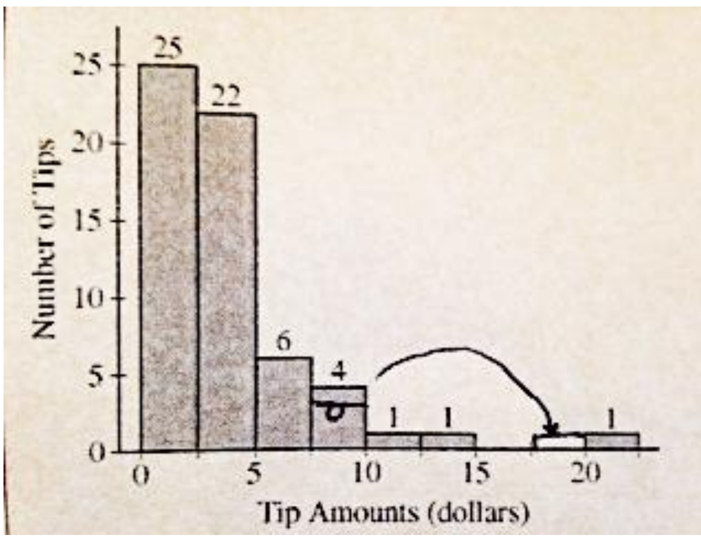


Figure 4.2. An example of reading within the graph drawn on the Tips Task.

During the interviews, three distinct types of transnumerative thinking (see Table 4.11) were provided: grouping a majority of the dataset between $0-$5, visualizing the $8 and $18 on the right-hand side of the data set, and realizing the median falls in the histogram bar containing 22 data points whether a tip was changed or not. All three transnumeration actions involved the ability to group data points and flexibly consider how changes in the shape of a distribution influence the median.

Respondents who used transnumerative thinking to notice the majority of the dataset was between $0-$5 suggested that the median should fall somewhere in this large-area which was a move in the right direction, but was not as advanced as transnumerative explanations. Respondents with this viewpoint were limited in describing the median's location, often stating that they would need to have the exact values of the dataset to be able to calculate the value of the median. Needing the exact values of the dataset to calculate the median is a true statement, but is still a limited view of what we know about the median. Four of the six respondents making statements about the median correctly answered the task.

A slightly more advanced transnumerative statement was suggesting that both $8 and $18 are on the same (right hand) side of the dataset. This statement suggested the ability to split the dataset into at least two groups and visualize where the median would lie in each of these groups. For example, Suzy described how she and her partner got to answer "E" by saying "the median does not change because one higher tip will not have an effect on the 25 and 22 lower tips." Unlike statements that suggested most of the data was between $0-$5, the five respondents who were able to make this statement were more confident in their answer choice, and all five respondents answered the task correctly. Respondents who used this more advanced transumerative thought process could organize groups of data with a median, keeping track of two different major ideas to solve the task.

Respondents who stated that a lot of points were in the 22 bar suggested knowledge that the median would fall in this group of the dataset. For example, Sean described his process of adjusting the histogram in the reflection section of the survey by saying "the median would stay in the range $2.50-$5 because many data points lie in that range." This transumerative thought process was considered more specific than realizing most of the dataset was between $0-$5 because it visualized the middle of the data set. Knowing the median falls in the 22 bar also suggested respondents could organize groups of data to the left and right of the median, however some of these respondents could visualize each histogram bar as individual data-points. For example, noticing that if six values in the 22-bar were $2.51, we know the median would be the average of the 30th and 31st number of $2.51. However, the mean could vary greatly still depending on the remaining values. The other 16 values of the 22-bar could all be $2.52 keeping a very low mean in comparison to having $4.99 for all 16 values. Breaking data-points apart within the 22-bar suggested a respondent could break groups of data apart with respect to measures of central tendency.

Nick was an Omega student who provided a good example of a transnumerative statement that involved advanced visualization of the dataset. When asked what information he knew from the graphical representations, Nick stated, "The median is between 2.5-5, the mean is somewhere on the right of 2.5, its probably around the $3, well I don't know its somewhere, somewhere above the median." Nick was able to state his visualization where the mean and median are within the histogram and utilize this information to describe the location of mean and median on the histogram.

In brief, evidence of visualizing the reorganization of the dataset with transnumerative thinking seemed to be key to responding correctly to the Tips Task. Twelve of the 16 respondents who showed evidence of using transnumerative thinking were completely correct in answering the Tips Task. In contrast, only 6% of the respondents who did not use transnumerative thinking were correct. Thus, visualizing "hidden" aspects in the dataset was a critical advantage when answering the Tips Task.

Specialized Knowledge of the Content in the Tips Task.

The Tips Task was an ideal topic for preservice teachers to demonstrate their Specialized Knowledge of the Content (SKC) given that the task discussed mean and median which were topics that preservice teachers understood, at least in some elementary manner. However, respondents really struggled to provide specific statements of how a teacher's knowledge of the Tips Task was unique in comparison with non-teaching professions. It was fairly common for respondents to discuss Common Knowledge of the Content (CKC) when trying to describe SKC. Four categories emerged during interviews from discussing how teacher knowledge was different than non-teaching professions, two CKC and two SKC. The first CKC category encompassed situations where preservice teachers stated some content that preservice teachers needed (e.g. "a teacher needs to know what a median is") without covering any comparison to other professions. The second CKC category included statements about how non-teachers should have the knowledge in this task, and therefore ignored any description of SKC by omission. Both of these kinds of statements avoided the purpose of the interview question, and only four of the nineteen (21%) preservice teachers who made these types of statements actually answered correctly. Two major categories of SKC emerged as well: that teachers should have knowledge of how changes in a data set affects measures of central tendency and a category that included all of the different, specific suggestions where transnumeration was discussed while highlighting SKC. Of the 16 respondents who made an SKC statement, eight answered the Tips Task correctly (50%).

Seven interviewees made statements that were coded into multiple knowledge subcategories below (see Table 4.12). Three of these interviewees made statements that were both in CKC categories, three interviewees made a CKC and SKC statement, and one interviewee made two coded statements in SKC categories. Only two of the seven (29%) interviewees who were coded for multiple knowledge subcategory statements answered the Tips Task correctly. Table 4.12 categorizes respondents' comments by CKC and SKC.

Table 4.12

Common and Specialized Knowledge of the Content exhibited by Preservice Teachers when completing the Tips Task.

	Alphas	Betas	Omegas
CKC			
-Stated a specific topic (e.g. median).	*Will, Jay,* Mia, *Dean, Cole*	*Sally* [2], *Jack, Chas, Kira* [2]	*Mary, Kelly* [2], *Julie* [2]
-Stated non-teachers should have this CKC.	Klara, *Dean* [2]	*Kadi*, Ben	Nick [2], *Julie* [2], *Kelly* [2], *Hope*
SKC			
-Teachers have SKC about changes to mean or median.	Suzy, *Lisa, Avis* [a]	*Ava, Amy, Sally* [2]	Kai, Kent, Jeremy, Myra [2]
-A specific suggestion using transnumeration with SKC.	Opal, *Neil* [a]	*Kira* [2], *Eve,* James	Nick [2], Myra [2]

Note. Some preservice teachers made multiple comments that fit topics, represented by number [2]. [a] represents a preservice teacher that was just agreeing with what another preservice teacher said during the interview. An *italicized name* indicates the student answered the Tips Task incorrectly, a regular font name indicates the student answered the Tips Task correctly.

Preservice teachers who answered with CKC statements failed to really answer the researcher's question(s) about SKC, promoting discussion about the knowledge differences between teachers and other working professionals. For example, when Jay asked about the knowledge teachers have that other professionals do not have, he only talked about the content he thought was needed to answer the Tips Task by stating "Knowing how histograms work. I guess like knowing what is shown on the histogram, like what the data on the graph is showing." Respondents in this category struggled to articulate that teachers have knowledge that other professionals do not. Respondents who stated non-teachers should have this CKC often were focused on the aspects of the task that are taught in middle and high school (e.g., a mean) and not the more complex aspects of this task, like analyzing Johnny's knowledge or considering where a mean or median would lie on the histogram. Of the eight respondents that stated non-teachers should know this material, only three (38%) answered correctly with one of those three also making SKC statements when responding to this interview question.

The comments that fell into the SCK category were divided into two subcategories. The first involved the difference between teachers and other professionals in knowing how mean and median are influenced by changes in a dataset. The second involved comments about how teachers use transnumeration where other professionals do not. Both of these types of responses showed an emerging conception of SKC. Five of the ten respondents (50%) who believed teachers know how changes to a dataset affect measures of central tendency correctly answered the Tips Task. Those preservice teachers who made specific transumerative statements with SKC correctly answered four out of seven times (57%).   It is not surprising that more interviewees who responded correctly were able to articulate how teacher knowledge is different than other professions. Interestingly, all five Omegas that made statements about SKC completed the Tips Task correctly. The overall struggle of all groups of preservice teachers to articulate SKC was surprising to the researcher considering the amount of classroom experiences preservice teachers have in schools. Most respondents believed teachers have different knowledge than other professions, but they struggled to give examples of such differences in the context of the Tips Task.

Pedagogical Content Knowledge in the Tips Task.

Quantifying PCK is a major challenge in education research. For this research study in particular, it was assumed that preservice teachers who lacked content knowledge in statistics would struggle to articulate pedagogical content knowledge in statistics. Statisticians generally accept that without content knowledge teachers cannot be affective. However a shift in thinking about what knowledge is critical to teaching has occurred in recent years. Moore explained that more attention is being given to pedagogical knowledge because this knowledge is the professional knowledge used daily (Ben-Zvi & Garfield, 2005). For the Tips task, some of the respondents who did not have the content knowledge that the median would not change when the $18 was changed to $8 still offered appropriate pedagogical suggestions on how to teach the content they understood. The following two sections about Knowledge of Content and Students and Knowledge of Content and Teaching provide respondents pedagogical suggestions.

Knowledge of Content and Students.

One of the interview questions asked respondents if a high school student would see anything different from the graphical representation than respondents themselves see. Twenty-two of the 32 interviewees (69%) considered their own knowledge of the statistics used to solve the Tips task to be the same as the knowledge of a high school student. Nineteen of these individuals (86%) justified their beliefs by talking about how the graph in the Tips task was basic to interpret (Table 4.13). The other three respondents actually made some kind of statement referring to their belief that high school students could use some sort of transumerative thinking when looking at this graphical representation.

Table 4.13

Respondents making KCS comments about the Tips Task

	Alphas	Betas	Omegas
No Difference			
-No difference, basic graph.	Klara, *Neil, Jay, Avis, Lisa, Will,* Mia, *Dean, Cole*	Ben, *Ava, Chas, Kira,* James, *Kadi*	*Mary,* Sean, Kai, Kent
-No difference, transnumeration reason.	Suzy	*Eve*	*Hope*
Different			
-Different, median would confuse students.	Opal	*Amy, Jack*	Carl [a], Nick, Myra
-Explicit graph difference		*Sally*	*Kelly*

Note. [a] was used to represent partially correct answers. An *italicized name* indicates the student answered the Tips Task incorrectly, a regular font name indicates the student answered the Tips Task correctly.

Dean was an Alpha who answered the Tips Task incorrectly but believed the content in the task was basic and the task should be easy for high school students to complete. During his interview, Dean said "I would say you could probably solve this in 7th or 8th grade math based on how relative the mean, median, and mode are." Of the eight respondents who believed there would be a difference between their knowledge and a high school student, four indicated that it would be difficult for students to understand the median remained the same with the data point change in the Tips Task. Myra was an Omega respondent who highlighted this difficulty stating:

> I think Johnny would probably be a pretty typical high school student. They just assume that mean and median are both measures of center so by having a lower tip mistakenly documented when it should have been higher, it should increase both of those when that's not really the case when looking at the data. So I think it would be important to highlight as a teacher and to really calculate out the mean and the median in this situation and try it with both cases to demonstrate and get that point across.

Perhaps the difficulty preservice teachers had in highlighting how a student would think about the Tips Task shows how hard it is to learn to think about student thinking during problem solving.

Knowledge of Content and Teaching.

Preservice teachers were asked two questions during the interview directed at understanding their Knowledge of Content and Teaching (KCT); one focused how they would teach this content to a twenty-student high school class and the other focused on how to teach this material to prepare students for the AP-Statistics exam. When asked about teaching this concept to a class, four different categories of answers emerged from the constant-comparison tables: (a) Data applications, (b) answer "E" (i.e., Have Johnny cut each histogram bar out and tape the bars together in order by length. Then have Johnny fold the entire length in half to show the median and discuss how this would change given the clerical mistake), (c) practice tip changes other than $18 to $8, and (d) change graphs (Table 4.14).

Table 4.14

Respondents KCT Statements about the Tips Task

	Alphas	Betas	Omegas
-Data applications	*Will,*	*Ava*, *Amy*, James	Nick, *Mya*, Kent
-Answer "E"	Klara	*Kira*	Sean, Kai, *Hope*
-Practice tip changes	Opal, *Neil*, *Cole*, *Jay*, Mia, *Dean*	*Jack*, Ben, *Kadi*	Carl, Jeremy, Myra
-Change graphs	Suzy, *Lisa*	*Sally*, *Chas*, *Eve*	*Julie*, *Kelly*

Note. An *italicized name* indicates the student answered the Tips Task incorrectly, a regular font name indicates the student answered the Tips Task correctly.

Preservice teachers were split relatively evenly within the four categories with eight respondents suggesting to use data applications, five respondents suggesting answer "E", twelve respondents suggesting to practice tip changes and seven respondents suggesting to change graphs. Suzy was the only one of seven preservice teachers that suggested change graphs (i.e., used transnumeration) and actually answered the Tips Task correctly on the survey. Perhaps this was because preservice teachers who made this suggestion struggled themselves to understand how the dataset was represented in the histogram. Five preservice teachers believed answer "E" would be how they would actually teach this content. This is an interesting finding because 15 preservice teachers answered only "E" on this task, leaving ten preservice teachers that believed they had other teaching techniques that would be better than answer "E." Recall, answer "E" was a specific teaching technique for visualizing a median in a histogram, suggested by an in-service AP-Statistics teacher. The ten interviewees that did not suggest they would use the answer "E" suggestion to teach stated more general teaching ideas (e.g. have students play around with another dataset). This response was surprising given that answer choice "E" *was playing around* with the dataset in a hands-on activity.

There was a noticeable difference in the manner preservice teachers who took AP-Statistics in high school explicitly mentioned statistics terminology and suggested teaching strategies with the exam in mind. Kira was a Beta respondent who took an AP-Statistics course in high school as a dual-credit option. When talking about her training for the exam, Kira said "We had a big emphasis on knowing the vocabulary and then at some parts knowing the equations, and a really big emphasis on being critical of graphs because not all graphs are organized correctly, or organized best for the information they are trying to put out. So you just have to be really critical." Even though Kira did not answer the Tips Task correctly, she provided specific details about what she looked for when answering questions. Eve was a freshman Beta student who earned a four on the AP-Statistics exam and a five on the AP-

Calculus exam. She had limited experiences with pedagogical coursework in college, but had substantial mathematics content background for a freshman due to the advanced placement courses she completed in high school. When asked about how she would prepare students for the AP exam, Eve said "Well my high school had a lot of take home packets and kind of taught the test and I'm a big supporter of practicing over and over to help so I think in order to prepare them have actual AP questions on your exam and stuff like that. I feel like if I didn't have those, I wouldn't have gotten a 4 on the exam or a 5 on the calculus exam. I think that practicing and having actual calculus problems on the test really made a difference."

Preservice teachers offered a variety of suggestions of how to teach the content from the Tips Task to help prepare students for the AP-Statistics exam (see Table 4.15). In comparison to how preservice teachers suggested they would teach material to a classroom of 20 students, 14 preservice teachers changed their statements significantly enough to be coded into a different KCT category for the AP students. Five preservice teachers suggested their technique would contain more attention to detail, and a fifth category was created called additional details. Of the five respondents coded for statements with additional details, only one actually answered the Tips Task correctly.

Table 4.15

Comparison of Changes in Response between KCT interview questions

	Data Application	Answer E	Practice Tip Changes	Change Graphs	Additional Details
-Data Application	*Will*, Avis, Nick, Kent		*Mya*, Nick	*Ava, Amy*	James
-Answer E		*Hope*, Kai	*Kira*	Klara, Sean	
-Practice Tip Changes		Myra	Jeremy, Ben, Carl, *Jack, Jay, Neil, Cole*	Opal, Mia	*Dean, Kadi*
-Change Graphs				Suzy, *Sally, Chas, Eve, Lisa*	*Kelly, Julie*

Note. The first column headings represent responses to teaching content to 20 students, the first row represents responses to how to prepare students for the AP-Stats exam. An *italicized name* indicates the student answered the Tips Task incorrectly, a regular font name indicates the student answered the Tips Task correctly.

An Omega student, Nick, built upon his first suggestion of having students play around with the dataset to see how the mean changes and the median remains the same by stating he would make the taught content more difficult to highlight certain statistical properties:

> I feel like that's a combination of the two: let them fiddle on their own and give them multiple examples. So they are not just working with the same data set and developing general rules based on a single problem. This one is skewed left, so if you are picking numbers, because big numbers are fun to move around you're going to change things. But what if you change like a $5 or $6 cost which is right near the center to a $0 tip, suddenly your data will change. But it's more-subtle than if you move one of these (pointing to low tips) up to here (pointing to high tips).

Tips Task Belief Responses.

Belief scores were compared between tasks. Survey factors, such as where each task appeared in the survey, may have influenced belief scores. The fact that the Tips Task appeared sixth out of eight on the survey may have influenced beliefs due to having already seen five previous tasks or to being tired of working through the survey.

Figure 4.3 shows the comparative difference between belief responses with the Tips Task and the mean score specific to Alpha, Beta, and Omega responses of all seven-belief tasks. Each belief question was broken-down by Alpha, Beta, and Omega groups. Scores above zero show that respondents agreed with that belief question with regards to the Tips Task more than the other seven tasks on average. Recall that belief responses were gathered on a 7-point Likert scale. The tallest bar (Beta respondents-Good) shows that on average Betas believed they were good at answering this task compared to the other tasks by almost a whole point.

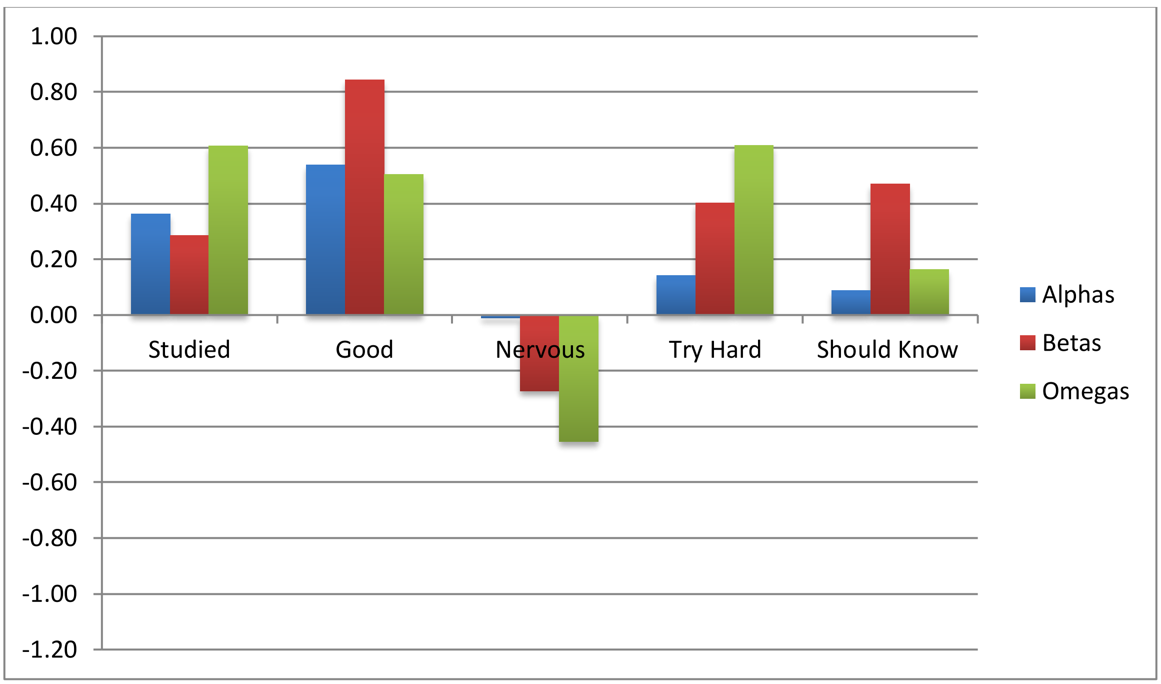


Figure 4.3. Comparisons of the mean response of belief questions to the Tips Task. Exact belief questions are as follows. Studied: This question is about a topic I have studied in a college class; Good: I am good at answering questions like this one; Nervous: I often feel nervous when I try to answer questions like this one; Try Hard: If I try hard, I can usually

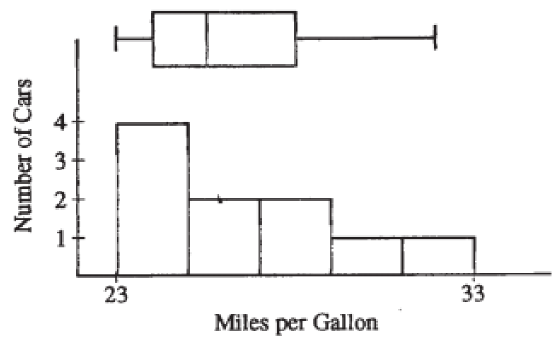
figure out questions like this one; Should Know: Secondary mathematics teachers should know how to answer this question.

As can be seen in Figure 4.3, on average, preservice teachers beliefs about the Tips Task were more favorable than the other six tasks the negatively-worded nervous belief question. This question, “I feel nervous about answering questions like this one,” was used as a check on belief responses as the lower the scores were, the more confident and better respondents believed they were at the task.   Thus, preservice teachers were actually less nervous on this task than the average task, or they felt better about answer this task compared to other six tasks.

Fuel Task Results

The Fuel Task (see Figure 4.4) was chosen for the survey because expert reviewers ranked it as the highest of the tasks that included transnumeration (a maximum-possible 3.0), and also because reviewers believed it would help answer research questions 1a and 1b, which focused on preservice teachers’ subject matter knowledge with graphical displays and transnumeration. The Fuel Task offered the opportunity to transnumerate between a boxplot, histogram, and table of descriptive statistics. Because the Fuel Task actually had two correct answers (“A” and “B”), respondents were required to transnumerate how a ratio comprised of measures of central tendency can be used to describe the skewness in a distribution.

Joe is a student of yours that is working on a problem with the following information: "A consumer organization was concerned that an automobile manufacturer was providing misleading information about average fuel efficiency by saying a new model of car gets 27 miles per gallon. The organization's researchers selected a random sample of 10 cars and assigned each to a random driver for 5,000 miles. The total fuel consumption for 5,000 miles was used to compute the mpg for each car. Below is a boxplot, histogram, and table that records the 10 sample values."



Minimum	Q1	Median	Q3	Maximum
23	24	25.5	28	32

Which of the following statement(s) could Joe make that shows some knowledge about skewness?

A. One way to describe the skewness of the data is the ratio $\frac{sample\ mean}{sample\ median}$. If the population is skewed to the right like above, the mean will be greater than the median, resulting in a large $\frac{sample\ mean}{sample\ median}$ ratio.

B. One way to describe the skewness of the data is the ratio $\frac{sample\ mean}{sample\ median}$. The closer the $\frac{sample\ mean}{sample\ median}$ ratio is to the value of "1," the closer the sample is to being symmetrical.

C. The mean is the vertical line inside the box in the boxplot pictured above.

D. Using the table, you could create a formula that describes skewness of the data. A formula that would measure skewness is $\frac{maximum}{Q3}$.

E. Using the table, you could create a formula that describes skewness of the data. A formula that would measure skewness is $\frac{maximum - minimum}{median}$.

Figure 4.4. The Fuel task. Answer choices "A" and "B" are both correct.

The Fuel Task was challenging for a variety of reasons: it involved multiple representations, it utilized a ratio of $\frac{sample\ mean}{sample\ median}$ in answer choices, and it involved skewness which researchers documented as one of the four major aspects of describing a distribution (Bakker & Gravemeijer; 2005). Both answer choices "A" and "B" required knowledge of how skewness within a distribution would affect the relationship between mean and median making the Fuel task even more challenging than had it only required transnumeration between graphical representations and the ratio $\frac{sample\ mean}{sample\ median}$. Answer "C" was designed as a distractor addressing the misunderstanding that the middle line of a boxplot represents the mean rather than the median of a distribution. Answers "D" and "E" were released examples from AP-Statistics of how *not* to measure skewness within a distribution. Preservice teachers did relatively well on this item, with almost 65% correctly selecting at least one of the correct answers "A" or "B" (Table 4.16).

Table 4.16

Results by group for Fuel Task

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)	Totals	Percentage
A	2	1	3	6	13.5%
B	5	4	6	15	40.5%
C	1	1	0	2	5.4%
D	1	0	0	1	2.7%
E	1	3	2	6	16.2%
A & B [a]	0	1	1	2	5.4%
B & C	1	0	0	1	2.7%
B & D	0	1	0	1	2.7%
Omitted	2	0	1	3	8.1%

Note. [a] denotes the best (i.e., completely correct) answer.

Answer "E", selected by 16% of respondents, was the most commonly selected incorrect answer. One possible reason that "E" was selected more than "D" was that, when asked during the interview, preservice teachers were unsure of what "Q" in "Q3" represented. Some preservice teachers were able to think about the limitation of trying to measure skewness with using the range divided by the median during interview discussions. Carl, recognized how unsuitable answers "D" and "E" were,

> I think if students give us an answer of "D" or "E" that gives us information that this student will need a lot of help for this content because D and E just mean nothing. So like if they had a formula like that, there would only be two functions. If not, they would have a very weird answer and they wouldn't know anything.

Only two respondents both Alphas, actually answered "C" even though five respondents made comments during the interviews about the middle-line of the boxplot being the "mean" while discussing their other answer choice. Respondents who discussed the middle-line of a boxplot believed they were not good at answering tasks like this one (see Table 4.17), averaging a response of 3.0 out of 7. There were also three respondents (Nick, Sally and Kelly) who stated bluntly that the median was the middle-line in a boxplot. These three respondents were a bit higher in confidence about answering questions like this one, averaging 4.3/7. Nick had advanced knowledge of the middle-line of a boxplot being the median as he used it to compare the mean: " Yes so 'C, is wrong probably. I don't have the data to calculate, but the mean is probably not the median of the boxplot based on the data I am looking at."

Table 4.17

Misunderstanding the Middle Line of a Boxplot

	Group	Answer	Good[a]	Comment
Lisa	Alpha	B & C	1	To Lisa, "C" was just a fact.
Chas	Beta	E	2	Agreed with Lisa that "C" was just a fact, and would have answered "C" and "E" if he took the survey again.
Mia	Alpha	B	1	Realized during the interview that the median is found on the boxplot. Stated: "I guess the box was the middle 50%"
Dean	Alpha	C	4	Noticed the median in the table of 25.5 was very similar to the middle-line in the boxplot. Thought mean and median should be about the same.
Sally	Beta	B	6	Stated "C" is wrong, the median is the middle-line of the boxplot, not the mean.
Ava	Beta	E	3	Implied the middle-line was a mean, stated "C only shows knowledge of mean."
Sean	Omega	B	3	Was not familiar with boxplots, also believed that the more area in a quartile the more data in that portion. Used the wording "to the left/right of the mean" to describe the middle-line of the boxplot.
Nick	Omega	A & B	5	Knew median was on boxplot, was comparing this line to where a mean line should be.
Kelly	Omega	B	2	Stated that the boxplot shows the median if you were not given the table with the five number summary of the distribution.

Note. Includes the seven respondents that commented directly to answer "C" or about the middle-line in a boxplot. [a] *Good* represents Likert scale responses to the belief question "I am good at answering questions like this one."

The most common answer choice for the Fuel task was "B," an answer that focuses both on understanding the $\frac{sample\ mean}{sample\ median}$ ratio and the skewness of the distribution. Including answers that had multiple responses, 46% of Alphas, 55% of Betas, and 54% of Omegas answered "B," which combined familiar topics from coursework preservice teachers should already have covered. Developing the mathematical idea of ratio begins in elementary school and is referenced frequently in coursework throughout middle school, high school, and even collegiate work. A distribution being symmetrical is a critical topic in the first undergraduate statistics course because distributions are the foundation for many calculations in a statistics course. In short, it is likely that a relatively high percentage of preservice teachers answered “B” because it combined familiar concepts that respondents could transnumerate information between.

Answer “A” was designed to evaluate how well respondents could envision a graphical representation containing measures of central tendency. Across all groups, respondents answered “A” only 22% of the time including the responses with multiple answers. While answer “A” began with the exact same sentence as answer “B,” the second sentence “If the population is skewed to the right like above, the mean will be greater than the median, resulting in a large $\frac{sample\ mean}{sample\ median}$ ratio” required respondents to think about what was happening in the distribution in a very different manner. The first concept that respondents often questioned was whether the distribution was actually skewed to the *right* and not the *left.* Of the 37 respondents, five incorrectly stated during interviews the distribution was skewed-left (James, Sean, Kent, Jeremy, & Myra) and therefore answer choice “A” was incorrect. Only three respondents stated correctly (Ava, Eve, & Kadi) that the direction of skewed-right was indeed correct.

Answer “A” was also challenging because it required respondents to compare mean and median on a graphical representation. On a histogram, knowing position of a mean-line in relation to position of a median-line is important to correctly interpret many dataset questions and is an example of transnumeration. Often a median turns out to be a much better measure of central tendency for contextual interpretations because a median, in contrast to a mean, is not pulled in the direction of outliers. Respondents who reasoned that in the Fuel task the mean would be greater than the median showed knowledge of a distribution's spread and not just the common measures of central tendency. A last likely reason why answer "A" was chosen less than "B" was the idea of a large $\frac{sample\ mean}{sample\ median}$ ratio. Respondents could have varying views on what constitutes large, limiting the number respondents who answered "A." Overall, the complexity of answer choice "A" was a hesitating factor for many respondents. An Omega respondent, Nick, answered both "A" and "B" as correct but described answer "B" as "more exact."

Fuel Task interview discussions that led to answer changes.

As previously noted, interviews provided opportunities for preservice teachers to rethink what they thought they knew about statistics and teaching mathematics. Kai and Kent were a good example of a pair of interviewees that developed knowledge throughout their discussion of the Fuel Task. Kent answered "B" by in his words "process of elimination," and Kai actually left the Fuel Task blank during the survey because she did not understand enough to respond. While answering interview questions, the conversation included:

> Kent – I thought "A" was backwards. I said because it said skewed to the right, the mean would be greater than the median, which I think it should be flipped.
>
> I – Why do you think it should be flipped?
>
> Kent – I don't remember, I just remember that's just what I thought when I read it. I can't remember exact what I was thinking it made sense when it said (if) the sample mean or the sample median is closer to one, then you would think being symmetrical made more sense. More than what "A" was saying.
>
> I – Why do you think that made more sense? Did you know what made it symmetrical?
>
> Kent – I think just because it was one, which seemed like a whole and good number. I don't know why, but symmetrical and one seemed like they fit together better than anything in "A."

> Kai – So I guess what it is saying is if the ratio is closer to one than that's saying that your sample mean is close to your median so it's close to the middle value. So whatever you have here on your upper quartile and lower quartile, or upper half and lower half, it's going to be closer to symmetrical because you have the average falls in the middle and your median falls where it is.

Kent and Kai both have gaps in knowledge with the concept of symmetry and how a $\frac{sample\ mean}{sample\ median}$ ratio is affected by a distribution, however there is substantial development throughout this conversation for Kai in discussing the mean and median and their relationship with symmetry. Specifically, Kai used statistical language (albeit not accurately as upper and lower quartiles are not each half of the dataset) in a comparative manner to assess the movement of the $\frac{sample\ mean}{sample\ median}$ ratio in a distribution.

Carl and Nick were a pair of Omega interviewees who were advanced in their ability to communicate their ideas with each other and work through tasks well. Both preservice teachers were at least partially correct on all tasks. Nick was so comfortable with the material that, at certain points in the interview, he actually began asking clarification questions to Carl:

> I – What math of statistics does a student need to know when answering this question?
>
> Nick - Knowing what skewness is a good step. Actually, Carl did you know what Q3 was? Because you didn't know what a box plot was.
>
> Carl - Yes I did. When I saw this I kind of figured it out. I figured that "Q" probably meant "quarter," so Q3 meant 75%.

Through questioning Carl, Nick began to distinguish his development as a teacher from the rest of his peers. First, Nick was astute enough to ask Carl a question to enhance Carl's learning opportunity during the interview. Secondly, Nick's question was thinking about student knowledge, or that the statistical content of knowing the five number summary and boxplot are related. Carl said he was never taught about boxplots in his schooling (perhaps because he was from a foreign country) and so for the Fuel task he spent time trying to extract information from the graphical representations. Nick linked quartiles being represented within a boxplot and asked Carl about his knowledge of quartiles, which was a really specific pedagogical content knowledge action. Nick was thinking about multiple aspects of a representation (demonstrating advanced CKC) and linking this to what another person knows about the content (demonstrating KCS).

Correct respondents' justifications for the Fuel Task.

Although 8 of 13 (62%) Alpha's answered the Fuel task partially or fully correct, interview results showed this group struggled with the content of this task. Two of the eight correct Alpha respondents (Neil & Jill) stated that they guessed to get the correct answer. When asked whether they agreed with the statement "I am good at answering questions like this one," the eight correct Alpha respondents averaged a 2.6 response out of 7, below the "somewhat disagree" score of three (see Table 4.18).

Table 4.18

Correct Alpha Student Justifications for Fuel Task

Respondent	Answer Choice(s)	Good[a]	Justification
Will	B	3	Will was not completely sure how to interpret graphical representations. He almost answered "A" in addition to "B."
Opal	B	5	Opal stated she knew that when mean=median, the data is symmetrical.
Neil	A	1	Neil guessed "A" because it was the longest and talked about mean and median.
Jill	B	4	Jill did not interview but her reflection said she guessed because she did not understand the concept of skewness well.
Suzy	A	2	Suzy liked "A," but also liked answer "E." She struggled to provide a definition of skewed.
Avis	B	4	Avis thought the closer the mean and median, the more accurate the data
Lisa	B & C	1	Lisa believed that if the ratio of sample mean and sample median was close to 1, it implied that the mean and median are similar.
Mia	B	1	Mia was confused on how to interpret the graphical representations. Skewness to Mia just meant "off."

Note. [a] Good column is the response to statement, "I am good at answering questions like this one" where 1 is strongly disagree and 7 is strongly agree.

Six of the 11 (64%) of Beta respondents were partially correct on the Fuel Task and an seventh (Kadi) was completely correct. Beta respondents who answered correctly averaged just above neutral 4.5/7 (see Table 4.19) in their responses to the belief question about whether or not they are good at answering this task. Interview data suggested that Betas had more knowledge about the statistical information in the Fuel task, although skewness was still difficult for many

Beta respondents. For example, James was able to describe skewness, but he struggled to use his definition to evaluate correct and incorrect answers. James said:

> Skewed is like if the data is stretched to fit a certain, nonstandard distribution. So in this sense it is a little skewed. It is favoring the front end of miles per gallon, or the lower end, so it's like skewed towards the lower end. That's why I got rid of "A", because it's not really skewed right, it is more skewed to the left.

James believed answer "A" was actually partially wrong because he stated the distribution was actually not skewed to the right, a common misunderstanding in the research literature (delMas, Garfield, & Ooms, 2005).

Table 4.19

Correct Beta Student Justifications for Fuel Task

Respondent	Answer Choice(s)	Good[a]	Justification
Jack	B	3	Jack answered “B” because he believed the mean/median ratio should by closer to “1” with a symmetrical distribution.
Ben	B	3	Ben could not remember exactly what skewness was, but guessed “B” and stated “I did not really know what was going on.”
Amy	B	4	Amy stated she had never seen the word skewness before in a statistics class. Crossed out answers C, D, and E.
Eve	A	6	Eve talked about skewed in a direction and the influence on a mean/median ratio.
James	B & D	4	James liked answer “B” because he understood how mean and median worked. James chose “D” because he incorrectly believed the distribution was not actually skewed right.
Taft	B	4	Taft did not interview but his reflection said choose "B" because he knew if the ratio of the sample mean and sample median is close to one the data would be symmetric.
Kadi	A & B	6	Kadi knew distribution was skewed right. Originally answered “B”, changed to “A and B” during interview.

Note. [a] Good column is the response to statement, “I am good at answering questions like this one” where 1 is strongly disagree and 7 is strongly agree.

Ten of the 13 (77%) Omega preservice teachers responded at least partially correct, including Nick who was completely correct. Omegas commonly indicated they had limited knowledge of skewness. Omegas’ usually reasoned through answer choices and arrived at “A” or “B” and indicated more confidence (3.8/7) then Alphas (2.6/7), but less than Betas (4.5/7) while answering the Fuel Task. Four of the seven correct responses of “A” came from Omegas indicating, that as a group, they had a deeper understanding of the direction of skewed distributions and ability to compare mean-line and median-line locations on a histogram.

Table 4.20

Correct Omega Student Justifications for Fuel Task

Respondent	Answer Choice(s)	Good[a]	Primary Justification
Mary	A	3	Mary struggled to articulate what skewness is, describing it as percent error.
Carl	B	4	Carl understood mean/median ratio well, but was not sure about how it represented skewness.
Nick	A & B	5	Nick believed "A" was true, but not as helpful to Jay as "B".
Sean	B	3	Sean believed incorrectly the histogram was skewed left, not right. Sean was more comfortable talking about the mean than median.
Kelly	B	2	Kelly was confused about skewness. Some knowledge of a boxplot (e.g., the median was the middle-line in a boxplot).
Kent	B	4	Kent used process of elimination with other answers to get to answer "B."
Mya	A	4	Mya did not interview but her reflection said she chose "A" because it was a mathematical ratio as well as a relationship with a skewed dataset.
Jeremy	A	6	Jeremy changed to "A" during the interview when he realized the graph was skewed right.
Chloe	B	3	Chloe did not interview but her reflection said she thought she understood skewness, but maybe not how to apply it.
Myra	B	4	Myra admitted to not knowing the formula for skew, but "B" made sense. Knew if skewed, mean does not equal median, if symmetric mean equals median.

Note. [a] Good column is the response to statement, "I am good at answering questions like this one" where 1 is strongly disagree and 7 is strongly agree.

Common Knowledge of the Content about graphical representations.

Interpreting the histogram, boxplot and table in the Fuel Task proved to be challenging for many respondents as their familiarity with each graphical representation was often limited in some manner. The task also used a skewed distribution intermingled with the ability to comprehend the $\frac{sample\ mean}{sample\ median}$ ratio, which were both abstract, transnumerative concepts. Using transnumeration begins with understanding parts of graphical representations, which will be discussed in the following section: histogram, boxplot and table.

Interpreting the histogram.

The histogram in the Fuel Task contained data for only ten total cars. The histogram was labeled appropriately and its dependent variable (i.e., number of cars) was explicitly depicted in scale through each histogram bar. There were challenging aspects of interpreting the histogram; one was interpreting its skewed-right nature (see Ch. 4 - *Interview discussions promoting change* for more details) and the scale of in independent variable (i.e., miles per gallon) being grouped by twos.

The concept that histogram bars have uniform length for a spatial representation using area was not always apparent to respondents. Kai was an Omega who answered four tasks completely correct (tied for the most of any respondent). Kai struggled to interpret the independent variable: "I think that maybe the *bar graph* would have been a little better if it had labels across there, but this is like ten spaces. I don't exactly know what those would be." From Kai's choice of terminology above, it's inferable that she did not understand differences between histograms and bar graphs when she called the Fuel Tasks' histogram and bar graph. She also could not break apart the independent variables scale by twos when she stated the graph would be better if it had *labels* to interpret the distribution.

Respondent's also struggled in interpreting the independent variable's range. The histogram's last bar ended at 33 because it scaled data into histogram bars that were two units each. However, the boxplot and table show the last data point being 32 miles per gallon. The variation between different graphical representations often was not interpreted correctly. Klara and Kira both answered "E," but had different understandings of the independent variables markings in the histogram:

> Klara – I remember thinking on how it started on the minimum but went past the maximum, because like it started at twenty-three which is where the minimum was, but then it went to thirty-three which is past the max. So like, I don't know. So they probably did it because they were going by twos, but it seems weird to me because you don't really know where thirty-three was because your max was thirty-two. So I don't know, I thought that was kind of weird.
>
> Kira – I am pretty sure that is this one because there was just one at thirty-two so they still had to go over here (to thirty-three), but that's a zero.
>
> Klara – Yes, I guess maybe. I don't know because how can you know there is one at thirty-three when your max is thirty-two. Yes, I don't know. It has nothing to do with the problem, but I still think it is weird. (laughing).

Klara's remark near the end of this transcript, "it has nothing to do with the problem" highlights a common trend among respondents of not knowing what information was important to extract from graphical representations. Furthermore, she did not see any relevance of being able to interpret the histogram's scaling, which limited ability to answer questions correctly.

The constant comparison table (Table 4.21) shows trends about respondent's knowledge of histograms. Perhaps the most intriguing trend was that respondents who were coded for a correct statement about histograms answered the Fuel task correctly nine out of ten times (90%).

In contrast, of the 15 respondents who made an incorrect statement concerning the histogram, only eight (53%) had a partially or fully correct answer for the Fuel Task. Two of the eight who answered correctly (Sally and Kelly) also made correct statements about the histogram. Additionally, four of the five respondents (80%) who discussed the context of cars within the histogram during interviews answered correctly. These findings suggest translating the histogram was critical to answering the Fuel Task correctly. It is likely that respondents who answered correctly had some understanding of how mean and median lines are influenced by skewness in a histogram, providing the background necessary to solve the task.

Table 4.21

Knowledge about Histograms in the Fuel Task

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
-Context cars & distribution [a]	Opal, Avis, *Cole*	Eve	Myra
Incorrect Histogram Statements			
-Independent variable confusion	Avis, *Klara*	Sally,	Mary, *Kai,*
-Called histogram a bar graph	*Lisa*		Kent, *Kai*
-Could not transnumerate between graphs	Opal, *Jay, Cole,*	Jack,	Mary, Sean, Kelly, *Julie, Kai*
Correct Histogram Statements			
-Understood 33 or 32 ending	*Dean*	Kira, Sally,	
-Explicitly correct statement about histogram	Neil, Suzy,	Amy, Kadi, James	Myra, Kelly

Note. An *italicized name* indicates the student answered the Fuel Task incorrectly, a regular font name indicates the student answered the Fuel Task correctly. [a] included respondents who made statements relating the task context with the graphical representation(s) of the distribution.

Interpreting the boxplot.

Respondents struggled to understand markings of a boxplots similar to previous research (Edwards, Özgün-Koca & Barr, 2017), perhaps because some aspects of a boxplot's structure are counterintuitive. Some respondents could interpret boxplots well, providing a large range of CKC across the respondents. Table 4.22 below includes statements about boxplots from both the Hiring Task and the Fuel Task because both tasks included a boxplot. This additional information about boxplots provides a better depiction of respondents' knowledge of the markings in a boxplot.

Table 4.22

Knowledge about Boxplots

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
Incorrect boxplot statements			
-Unfamiliar or generally incorrect boxplot statements	Opal, Neil, Mia, Lisa, Will, *Cole*, *Jay*, *Klara*,	Jack, James, *Chas*,	Sean, Kelly, Carl, Kent, Mary, *Hope*, *Julie*,
-Quartile confusion based on size of box.		*Ava*	Mary, Sean,
Correct boxplot statements			
-Knew median, quartiles, or IQR locations.	Avis, *Jay*, *Cole*,	Sally, Eve, Ben, Kadi[a], *Ava*, *Kira*	Kelly, Carl, Myra, Nick[a] Kai
-Skewed boxplot meant different mean and median		*Ava*	Nick[a], Mya[b]
-Understood why boxplot ended at 32	*Dean*	Sally, *Kira*	
-Linked table to boxplot (transnumeration)		Kadi[a], Sally Ben, *Kira*	*Julie*, Sean, Myra
-Drew lines from boxplot to histogram (transnumeration)	*Klara*		
Belief Statements			
-Boxplots help visualize		Sally, Kadi[a]	Jeremy,
-Table makes boxplot irrelevant	Suzy,	Amy	*Julie*, *Kai*

Note. An *italicized name* indicates the student answered the Fuel Task incorrectly, a regular font name indicates the student answered the Fuel Task correctly. [a] Answered completely correct. [b] Survey reflection was source of code.

Respondents who felt like the boxplot was unnecessary because they were also given a table were not completely wrong; the table gives a five-number summary with the vertical markings in the boxplot. However, the four respondents who said a boxplot was unnecessary did not know enough about boxplots to interpret a distribution. When discussing boxplots, Suzy stated it had "been a very, very long time since I have seen one." Amy indicated that each number in the five-number summary was on the boxplot, but stated "I don't know what you are really looking for." Julie relied on her partner Kelly to take the lead during boxplot discussion even though Kelly made multiple wrong statements about boxplots. Kai was probably the most knowledgeable because she had recently observed a middle school lesson about boxplots, but she really struggled to transnumerate data between the boxplot, table and histogram and did not have a conception of the $\frac{sample\ mean}{sample\ median}$ ratio.

Nineteen of the 37 respondents either stated that they did not know how to interpret a boxplot, had never seen a boxplot before, or made multiple incorrect statements about the markings on a boxplot. Interestingly, 12 of these respondents (63%) actually answered part of the Fuel Task correctly. Of the 21 respondents who made correct statements about boxplots, only 11 (52%) actually answered the Fuel Task correctly, which suggests that the lack of knowledge respondents had about boxplots did not keep them from figuring out enough information from the boxplot to answer the Fuel Task correctly. Perhaps this was because many respondents used the histogram to visualize the relationship between the mean and median (see Figure 4.6 for an example) in a skewed distribution.

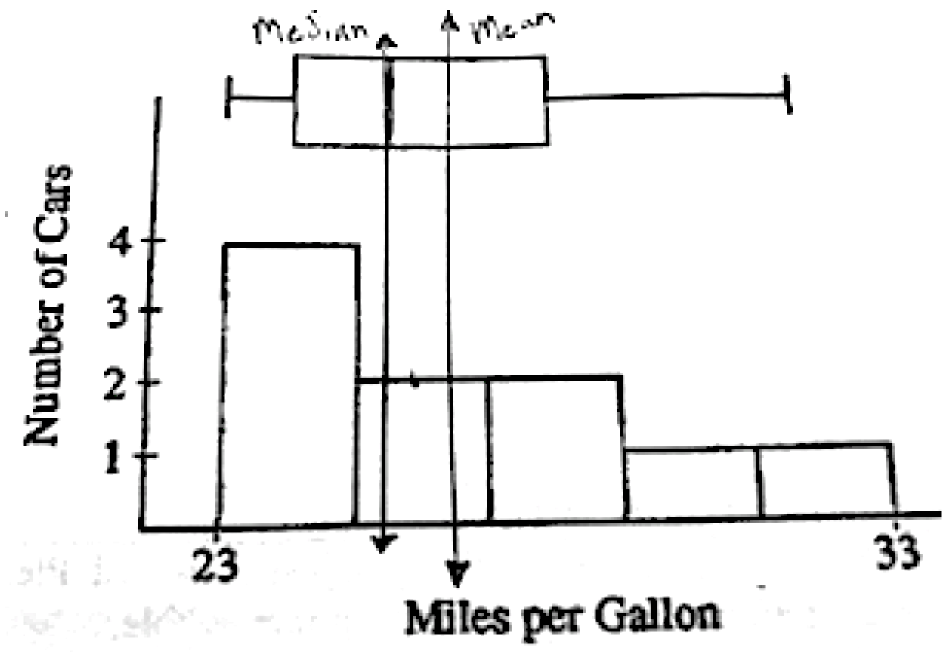


Figure 4.6. An example of a skewed-right distribution with mean and median lines.

The two respondents (Kadi and Nick) who answered the Fuel Task completely correct were coded as making multiple correct statements about boxplots. This suggests that when respondents could state multiple attributes of a boxplot, they had enough CKC to interpret the accuracy of answer choices A and B. It's important to note both of these respondents took statistics in high school where they likely had frequent exposure to boxplots.

A misunderstanding about boxplots.

One unique aspect of the Fuel Task was that the boxplot and histogram shared the same axis for the independent variable, which was scaled by miles per gallon of gasoline. This created some interesting comparison opportunities with the material presented on the graphical representation. For example, histogram bars *gain* visual area represented more data-points, whereas boxplot quartiles actually *lose* visual area represented more data-points. Other researchers found boxplot quartiles losing visual area was counterintuitive and often confusing for novice boxplot readers (Lem, Onghana, Verschaffel, & Van Dooren, 2013). Sean was an Omega, a graduate student who was completely correct on four of the survey tasks and partially correct on the remainder. Sean struggled with the area representation of the boxplot during the Fuel Task,

> So because I can see from the boxplot that the area to the left of the mean is a smaller area (than) to the right of the mean. So that makes me think there are more data points to the right of the mean, which would mean that the median is also to the right of the mean, meaning its larger. So I think I got that information from the boxplot.

Sean's graduate-level mathematical and reasoning capabilities often helped him work through tasks to correct answers, but some of the less-intuitive statistical concepts were a struggle. More area in a graphical representation usually means a larger quantity, but not with boxplots.

Interpreting the table.

The table of values in the Fuel Task was a simplistic graphical representation as long as respondents had experience with the content headings presented. In conversations about the task, the table received the least attention. This could have been because the table was viewed as a representation that just stated information bluntly and respondents rarely struggled with interpreting the table. The only incorrect statement based on the table involved the fact that some respondents did not know what Q1 or Q3 stood for in the table (see Table 4.23 below). The respondents specifically said they did not know what Q1 or Q3 meant but many others incorrectly attempted to interpret the meaning. For example, Carl stated: "When I saw this [the table] I kind of figured it out. I figured that Q probably mean quarter so Q3 meant 75%." Carl could not interpret the boxplot initially because he had not seen it before, but his responses show that a respondent without knowledge of a boxplot might be able to make connections to the five-number summary in the table.

Table 4.23

Knowledge about the Table in the Fuel Task

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
Incorrect Table Statements			
-Did not know Q1/Q3	Neil	*Chas*	Carl
Correct Table Statements			
-Used table to solve	Neil	James,	
-Table gave useful numbers	Suzy, Lisa		Madi T, *Hope*
-Table tells median, max, min	Opal, Neil	Kadi, Ben, *Ava*,	*Kelly*
-Knew Q1/Q3	Lisa,	Sally, Kadi, Amy, *Ava*,	Sean, *Hope*,
-Apply table numbers to formulas	*Cole*,		Sean
Transnumeration Statements			
-Table links to boxplot parts	Lisa, *Jay*, *Cole*, *Klara*,	Sally, Kadi, Amy, *Ava*	Sean, Myra, *Hope*, *Julie*, *Kelly*, *Kai*,
-Boxplot make table irrelevant		*Ava*	Myra, *Julie*,

Note. An *italicized name* indicates the student answered the Fuel Task incorrectly, a regular font name indicates the student answered the Fuel Task correctly.

Of the 15 respondents who made explicit correct statements that involved the table, 11 (73%) completed the task correctly. Because the table was the third representation from top-to-bottom displayed in the Fuel Task, when preservice teachers discussed material in the table often it showed additional attention to the task. However, the table was limited in the information to answer the task correctly. Visualizing the relationship of mean to median on a distribution with the ratio $\frac{sample\ mean}{sample\ median}$, brought challenges that demanded transnumeration, which was difficult to utilize in the table. Without including a sixth quantity in the table summary, a sample mean, the information given provided was incomplete to get to answers “A” or “B.”

Answers “D” and “E” could be calculated using only the table values to answer the question. Only two respondents, James and Jay, provided the correct response of “D.” James interestingly spent more time discussing the table than the other graphical representations, but also chose answer “B.” Jay stated that he answered “D” as an “educated guess,” while thinking about the task as links between representations. Except for Kira, the six respondents choosing “E” as an answer (Klara, Ava, Chas, Kira, Julie and Hope) were coded as mentioning the table during the interview. This suggests respondents choosing “E” may have limited knowledge of how to interpret statistical representations (i.e a histogram, boxplot) and thus focused on the straight-forward information provided in the table instead.

Fourteen individuals stated during interviews that the table’s information linked directly to the boxplot. Only three of these 14 were able to state that the boxplot showed more information about the distribution then the table, affirming how uncomfortable respondents were with understanding how a boxplot shows the spread of a distribution visually. Respondents who commented on the connection between the table and boxplot had a markedly worse average response rate then a typical respondent, with only five of the 14 (36%) respondents at least partially correct on the Fuel Task. Recall, 70% of respondents were at least partially correct on the Fuel Task overall. Because answers “D” and “E” could be completed from the descriptors in the table, they appealed to respondents without CKC of histograms or boxplots. In contrast the correct answers “A” and “B” were not dependent on using the tables descriptors as much as transnumeration visualizing the distribution.

Transnumeration between graphical representations.

The Fuel Task created unique opportunities for respondents to use transnumerative thinking across different representations because it had three graphical representations that all described the same data set. Respondents had the opportunity to link attributes in three different graphical representations together, seeing important connections about the dataset that were described in each representation. Sean, an Omega who was a mathematics graduate student, was interesting as he had a lot of mathematical knowledge but limited statistical and pedagogical knowledge. When Sean was asked how he would teach the material in the Fuel Task to a class of 20 students, he suggested using multiple representations with the same dataset as a teaching strategy:

> Sean – Probably drawing connections between the three different things (graphical representations) that were shown.
>
> I – Ok, why do you think that's important?
>
> Sean – (pause) I think it's good statistical knowledge as a whole, so not just to this particular problem (pause). Maybe related to this problem, once you drew those connections you could figure out ways to measure the skewness. I decided that the data was skewed by looking at the boxplot, but say that you're able to make that leap with any one of these things (graphical representations), maybe if you translate the information into another picture, another form, of representing the data like the table or the histogram, maybe that can tell you more information about the skewness or about how that works.
>
> I – When you say that word "information," are you talking about concepts or ways to solve this problem? (Sean cut me off)
>
> Sean – Maybe it could give you more ways to measure the skewness.

Sean likely developed confidence through his experiences of taking graduate level mathematics courses, and believed he could interpret attributes in the graphical representations even without necessary statistical literacy (e.g., Sean thought the boxplots middle-line was the mean, not the median). Sean's confidence was likely supported by previous near-transnumerative experiences between graphical representations where he acquired statistical knowledge.

Markings that supported transnumerative thinking.

The researcher expected respondents to draw markings on the Fuel Task to communicate knowledge between different representations. Evidence from the survey showed only eight of the 37 (22%) drawing on their survey (see Table 4.24). The most common markings made were writing in or referring to the scaling of the histogram by two units in each histogram-bar, with four respondents explicitly writing in interval numbers.

Table 4.24

Markings Used to Answer the Fuel Task

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
Individual markings			
-Wrote histogram bar height			Mya [b]
-Independent scaled by twos	Lisa, *Dean,* *Klara,*		Chloe [b], Carl, Nick
Boxplot to histogram			
-Median line only			Nick
-Q1, median line, and Q3	*Klara*		*Hope* [a]

Note. [a] Marked her graphical representations during the interview. [b] Only took survey. An *italicized name* indicates the student answered the Fuel Task incorrectly, a regular font name indicates the student answered the Fuel Task correctly.

In addition to scaling the histogram, respondents sometimes labeled the start and finish points of each histogram bar to show the independent variables were scaled by twos. Dean was an Alpha respondent who answered “C,” but marked numbers on his survey showing the independent variables scale of two (see Figure 4.7). Dean said that he added this information to the histogram in an attempt to determine where the median of 25.5 was located.

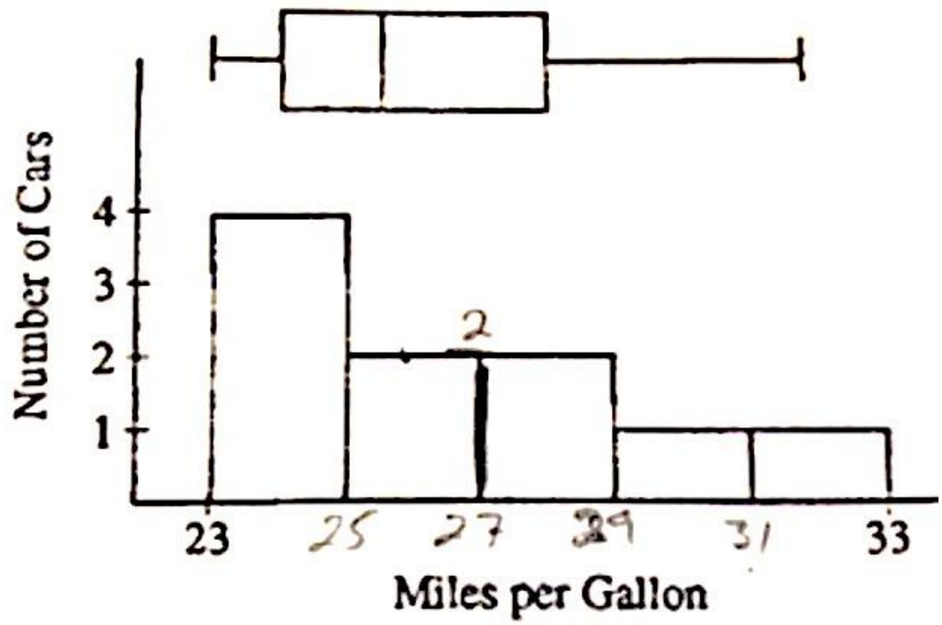
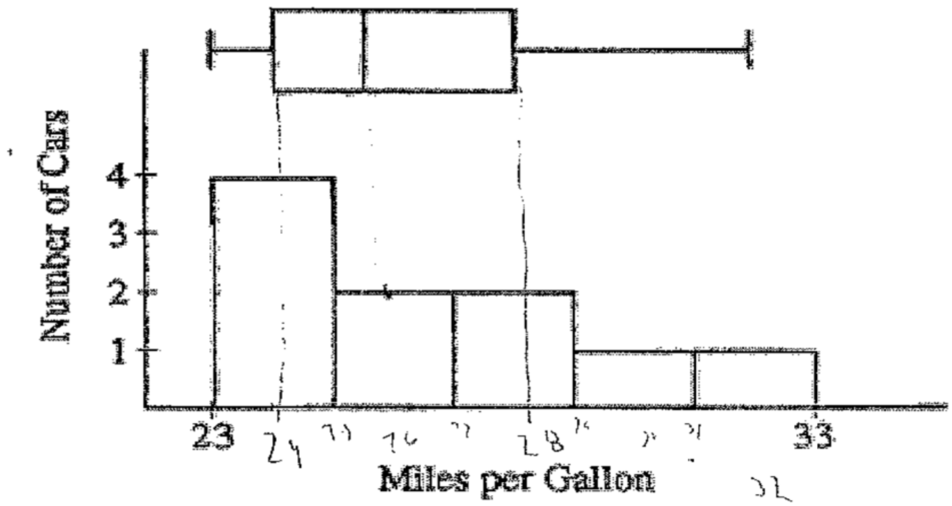


Figure 4.7. Dean's markings of histogram lengths while taking survey.

It was rare for respondents to show transumerative thinking between all three graphical representations, but in some cases there was written evidence of this kind of thinking. Hope (an Omega) wrote in independent variable scale numbers during the interview in an attempt to find the median for the dataset on the histogram like Dean, or in her words to "make sense of where the median sits." This statement provided evidence that Hope did not interpret information about the independent variable from both the boxplot and histogram. When respondents realized data was linked between representations, a common application was to use transnumerative thinking between the graphical representations. Klara made the most markings between the boxplot and histogram, drawing vertical lines down from the boxplot onto the histogram including linking the median and circling the median on the table (see Figure 4.8). Although Klara did not get this task correct, there was a clear effort to transnumerate the knowledge between each of the graphical representations for a more cohesive view of the distribution.



Minimum	Q1	Median	Q3	Maximum
23	24	25.5	28	32

Figure 4.8. Klara's transnumeration markings on the Fuel Task.

Nick was perhaps the most advanced respondent in knowledge across all four categories. Nick tied all respondents for both the most tasks at least partially correct with seven and the most tasks completely correct with five, but what distinguished him even more was his ability to talk about his CKC during the interview. Nick often moved quickly during interview discussions, highlighting information he extracted from graphical representations and used from each task. Nick's markings in the Fuel Task were simple, but showed interaction between the different representations (see Figure 4.9).

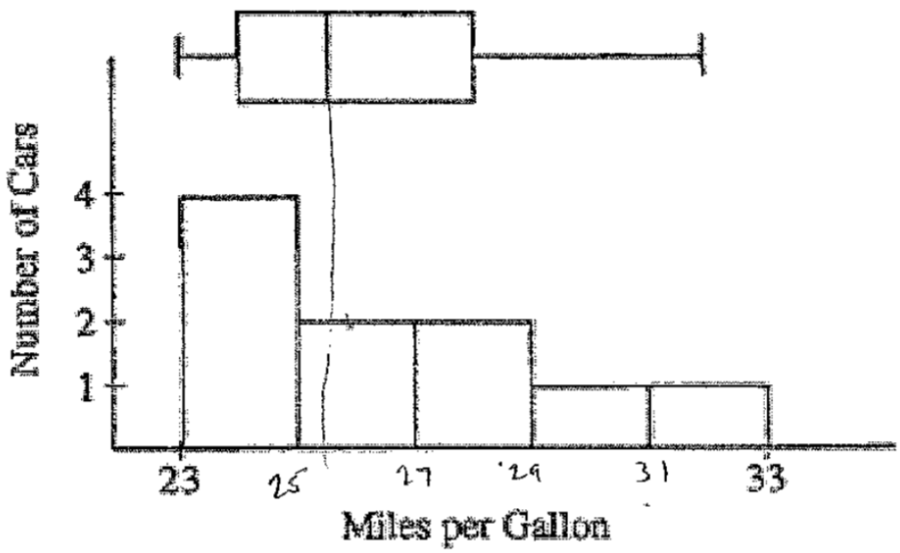


Figure 4.9. Nick's transnumeration markings on the Fuel Task.

When Nick was asked "what information do you know from these graphical representations?" he said:

> I know the range. I know what percentage of the data is in each zone. I know median. I don't know the mean because I don't know all of the data. I can't calculate the mean without more (information). I know that the mean will be greater than the median.

Nick provided a clear list of statistical descriptors along with information about which descriptors (i.e., a median) were known and which (i.e., a mean) could not be found without knowing the exact data-points.

Transnumeration involving mean and median on representations.

Respondents' use of transnumeration with markings was considered an easier cognitive operation compared to when respondents used transnumeration with the ratio $\frac{sample\ mean}{sample\ median}$. Perhaps this was because a flexible, visual understanding of mean and median was already difficult for respondents and comparisons of these units *as a ratio* was even more difficult.

Table 4.25 displays preservice teachers comments about the relationship between mean and median.

Table 4.25

Transnumeration involving the Relationship between Mean and Median

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
Incorrect Statements			
-About $\frac{sample\ mean}{sample\ median}$	*Dean, Cole*		Sean, *Kai*
-About skewness	Jill, *Bell*	Ben, Amy	Mary, Chloe[a], Kelly, *Julie, Hope*
Correct Statements			
-If $\frac{sample\ mean}{sample\ median}$ =1, distribution is symmetric	Will, Opal, Avis, Lisa	Sally, Jack, Eve, James, Taft[a], Kadi	Carl, Nick, Sean, Myra
-On skewed graphs, mean influenced more than median	Neil, Avis	Eve, Taft[a], Kadi, *Ava*	Nick, Kent, Mya

Note. [a] This respondent only took the survey. An *italicized name* indicates the student answered the Fuel Task incorrectly, a regular font name indicates the student answered the Fuel Task correctly.

Respondents' statements about the relationship between mean and median were categorized as incorrect or correct. The two incorrect statement subcategories focused on either misinterpreting the $\frac{sample\ mean}{sample\ median}$ ratio or stating a limitation in knowledge about skewness. Thirteen respondents made one of these statements, with seven (54%) responding to the task correctly. It was not surprising that three of the four respondents who did not recognize the ratio $\frac{sample\ mean}{sample\ median}$ answered incorrectly because both "A" and "B" used this ratio explicitly in the answer choice.

Perhaps more interesting was how successful respondents were when they made correct statements about the relationship of mean and median. Two subcategories emerged within correct statements, the first included statements about how, when the mean and median are equal, a distribution is symmetrical. All 14 respondents who indicated that when the mean and median are equal the distribution is symmetrical were at least partially correct! This suggests that understanding the ratio of $\frac{sample\ mean}{sample\ median}$ was less of an issue as long as a concept of symmetry and mean and median location were present. The other correct subcategory focused on how a mean was influenced more by outliers or pulled in a skewed direction, which in this case meant the mean was greater than the median. Of the nine respondents that were coded for this statement, eight (89%) were correct. This showed when respondents understood the relationship of mean and median on a graphical representation, there was a high likelihood of being correct on the Fuel Task. There was some overlap between the two correct subcategories, with five respondents coded for statements in both subgroups.

Only eight respondents selected “A” as either part or all of their response to the Fuel Task. Of those eight respondents, only three (37.5%) failed to make statements during interviews coded as correct statements in Table 4.25. The two respondents who answered completely correct (Nick and Kadi) made statements that were categorized into both categories. This suggests that in order to understand all of the intricacies of the correct answer choices in the Fuel Task, respondents needed knowledge about symmetric or skewed distributions and how measures of central tendency align within distributions.

Respondents understood the importance of transnumeration even when they answered the Fuel Task incorrectly. For example, Omega group respondents Julie and Kelly said the following about transnumerating across representations:

Julie – You can see the data and compare the two (graphical representations) and see where they are similar and different.

Kelly – I think it also highlights that you can see how the boxplot is shifted to the right. So then you notice more with the histogram, oh there is more of a correlation towards twenty-three so I guess that would highlight that more.

Julie – I think it also gives you the opportunity to show different ways of representing the data too and that they may look very different, but they are telling the same thing.

I – Do any differences in particular jump out to you?

Kelly – Well I think with mean and median looking at the histogram, finding the median is more difficult, but then using the histogram (meant boxplot).

Julie – You mean using the graph (histogram), its more difficult.

Kelly – Sorry histogram and boxplot, sorry I switched those two, but the boxplot shows that (the median) if you were not given the table.

Kelly and Julie are limited in being able to explain their knowledge, and have limited CKC, but they do see the importance of providing opportunities to students to view attributes of datasets. The next section presents results about respondents SKC during the Fuel Task.

Specialized Knowledge of the Content in the Fuel Task.

Specialized Knowledge of the Content (SKC) was evaluated by asking respondents what knowledge teachers have with this material that other working professionals do not have. As was the case with the Tips Task, identifying knowledge about the Fuel Task that teachers need beyond what other professionals was difficult for interviewees. Respondents mentioned different content topics or brought up CKC most often, similar to the Tips Task. Only six individuals were coded for SKC responses (see Table 4.26). Five of the six respondents (83%) answered at least one correct answer.

Table 4.26

Specialized Knowledge of the Content from preservice teachers from the Fuel Task.

	Alphas	Betas	Omegas
CKC - Generic			
-Teachers need all material	*Jay, Dean*		
-Know context of graph.	Will, Mia, *Klara, Cole*		*Hope*
-Know equations / formulas	*Klara*	Sally,	
-Know the Vocab		James, *Kira*	
CKC – Specific Stats			
-Know median / quartiles	Will, Avis, Lisa	Sally, Ben, Amy	Nick, *Hope*
-Know skewed data	Suzy, Lisa, *Cole, Klara*	Jeremy, *Ava*	Nick, Kelly, Myra
-Know boxplot or histogram	*Chas*	Jack, Eve	Mary, Kelly, Myra
CKC - Transnumeration			
-Know how boxplot and histogram relate		Jack, Ben, Jeremy	Mary, Kent, *Hope*
-Know relationship of mean-median with a dataset	Opal		Myra
SKC			
-Know how to explain specific statistics ideas	Neil	James	Mary
-Task is unique knowledge of teachers		Kadi	Nick

Note. An *italicized name* indicates the student answered the Fuel Task incorrectly, a regular font name indicates the student answered the Fuel Task correctly.

CKC was coded into three categories to highlight the differences in respondents' comments. The four CKC-generic subcategories were teachers stating they need to know this material, know the graph or context of the graph, know the equation or formulas, and know the vocabulary. Although multiple respondents mentioned these subcategories, most of the comments made were general rather than specific to statistics. For example, Kira just said teachers need to "do the vocab", as opposed to talking about critical, specific vocabulary terms in this task like a sample median or skewed distribution. Of the nine individuals who were categorized as making a generic statistical statement, five responded incorrectly on the task.

The majority of respondents were coded for making a CKC-specific stats statements categorized into three specific subcategories: the median or quartiles, skewed data, or boxplot or histogram. Of the 21 respondents that made CKC-specific stats statements, 16 (76%) answered at least partially correct which was a substantially higher correct response rate than the respondents who made CKC-generic comments. Four of the incorrect respondents (Cole, Klara, Sally, Hope) were also coded for CKC-generic statements, meaning their language reverted to simplistic, un-statistical explanations during conversations in places where statistical language was helpful. CKC-generic statements were most often made by Alphas(6) as compared to Betas(3) and Omegas(1) suggesting experience with statistics and pedagogy made a major difference in the quality of SKC responses.

The third subcategory called CKC-transnumeration included statements that used transnumeration between a boxplot and histogram *or* knowing the relationship of mean and median within a dataset. Eight respondents, including seven who answered the Fuel Task correctly, suggested transnumeration as knowledge teachers have that other professionals do not. Four of these respondents made CKC specific category statements. The one Omega in this subcategory (Hope) made a CKC-generic statement. Although Hope answered the Fuel Task incorrectly, she was particularly focused throughout her interview on pedagogical aspects of mathematics. Although she admitted to not knowing what a boxplot was, or how skewness worked in a distribution, she made statements about how teachers need to know this content when they are thinking about students and how to teach this material. Her interview suggested she knew the importance of both content and pedagogical knowledge for a teacher.

There were only five respondents whose statements were coded into the Specialized Knowledge of the Content (SKC) category into two different subcategories: knowing how to explain specific statistical ideas like a median, skewness or different representations strengths *and* stating teachers will know this material when other professionals will not. Kadi and Nick, the only two respondents who were completely correct on the Fuel Task, stated that teachers will know this material when other professions will not. Perhaps as interesting was that only five of the 32 interviewees (16%) were coded for a SKC comment on the Fuel Task, all of whom responded at least partially correct.

The next two sections describe comments made in the interviews about KCS and KCT. KCS conversations typically were responses to a comparison question where respondents were asked to think about the difference between a high school student's knowledge of the content and their own knowledge. KCT comments focused both on teaching a class of students and helping students to prepare for an AP-Statistics exam.

Knowledge of content and students.

Although the Fuel Task was written in part to assess Knowledge of Content and Students (KCS), respondents struggled to articulate meaningful statements in this knowledge category. Survey and interview results showed that the content in the Fuel Task was very difficult for respondents. The Fuel Task also had multiple correct answers, which meant that respondents had to reason through each answer choice. Interview results showed respondents struggled to articulate what a high school student would see different in this graphical representation. Responses were coded into two broad categories of there being a difference in knowledge or not, and then by more specific subcategories (see Table 4.27).

Table 4.27

Knowledge of Content and Students from preservice teachers from the Fuel Task

	Alphas	Betas	Omegas
No Difference			
-No difference	Suzy, *Dean,*	*Ava*	*Kai*
-I am confused	*Jay*		Kent
Difference			
-Not sure how		Jack	
-Skewed distribution would be hard	Lisa, *Klara*	Amy, Kadi	Nick, Sean, Myra, Kelly
-Transnumeration would be hard	Neil	Sally, *Kira*	Nick
-Boxplot would be hard	Opal, Avis	Ben, Eve	*Hope*
-Context would be hard	*Cole*	James, Kadi	Myra

Note. An *italicized name* indicates the student answered the Fuel Task incorrectly, a regular font name indicates the student answered the Fuel Task correctly.

Respondents who stated that there was no difference between what a high school student and a teacher would see on the Fuel Task did proportionately worse on the Fuel Task than those who saw a difference with two out of six (33%) getting the question partially correct and the rest not able to answer any part of the task correctly. This was not that surprising given that a respondent who could not think about differences between students and themselves as teachers, typically did not think they had a good grasp of the content presented in the Fuel Task. Respondents who pointed out differences between themselves and students did better than those who did not, with 15 of the 19 (79%) answering at least partially correct. Only three preservice teachers were coded for making multiple statements about the differences between students and themselves, with two of those being Nick and Kadi (the preservice teachers who answered completely correct). Nick was a good example of a respondent who believed his future students would have different knowledge about this problem than he does. He demonstrated KCS when saying, for example, “I can definitely see how they might think it’s skewed to the left because the left is so much bigger,” which is a common statistical misconception.

When analyzing KCS data holistically, the most frequent subcategories mentioned involved (i) the idea of skewed distributions being difficult and (ii) boxplots being difficult to interpret. Previously described CKC subcategory responses aligned to these KCS findings, as nine students were coded for making incorrect statements about skewness (see Table 4.25) and 19 respondents stated that they were unsure, or did not know how to interpret all of a boxplot (see Table 4.22). Where respondents had limited CKC, they often suggested that their students would have similar limitations with the content.

Knowledge of content and teaching.

Respondents’ suggestions for teaching the concepts in the Fuel Task to a class of 20 students were coded into three categories: traditional pedagogical, reform pedagogical and transnumeration (Table 4.28). Traditional pedagogical suggestions focused on two areas: reiterating the content *and* lecturing or explaining what the content means to students. Of the 12 respondent statements who were coded into the traditional pedagogical category, only six (50%) responded to the Fuel Task at least partially correct.

Table 4.28

Knowledge of Content and Teaching from preservice teachers from the Fuel Task

	Alphas	Betas	Omegas
Traditional Pedagogical			
-Revisit content again	Will, *Dean*	Jack, *Chas,*	*Hope*
-Lecture or explain	Lisa, *Klara*	Amy, Kadi, *Kira*	Mary, *Julie*
Reform Pedagogical			
-Group work	Mia,		Kent
-Ask questions	Neil		*Hope*
-Students make graphs		James	Nick, Myra
-Real-life application	Will, Avis	Sally, Eve	Jeremy
Transnumeration			
-Compare multiple graphs representing distributions	Opal, Neil, *Jay*		Sean, Kelly, *Kai*
-Use multiple skewed distribution	Suzy, *Cole*	Ben, *Ava*	Nick, Carl

Note. An *italicized name* indicates the student answered the Fuel Task incorrectly, a regular font name indicates the student answered the Fuel Task correctly.

The reform pedagogical category was broken into four different subcategories: group-work, ask questions, students make graphs and real-life applications. Of the 12 respondents who were coded into one of these categories, 11 (92%) responded correctly. The most common suggested direction in the reform pedagogy category was using real-life applications as part of instruction. This pedagogical tool was also one of four suggestions that emerged in the Tips Task.

The last group of KCT subcategories suggested to use transnumeration to teach this concept by comparing multiple graphical representations *or* giving students multiple examples of skewed distributions. Of the 12 respondents that were coded for a transnumeration focused suggestion, eight (67%) responded to the Fuel Task correctly. Opal was an Alpha coded for suggesting to teach by comparing multiple graphs representing distributions:

Opal - I would say I would want my students to write sets of data in order to understand skew versus symmetry because if it is sample mean over sample medium there are three possibilities. The mean is greater than the median, they're equal, or the median is greater than the mean. I would want them to see data that represents all those and what all three of these would look like to be able to get more meaning out of what skew is. I know that skew is either symmetrical in the middle or to the left or right and so I want them to see all three of those to understand the context of what it would mean if you had the data skewed or symmetrical.

I - Okay, why do you think that would be helpful?

Opal - I don't know, just telling me this is skewed doesn't really mean anything to me unless I can see what it is skewed compared to.

I - Okay, when you say that word compared to, do you mean something else in general?

Opal - Like it being symmetrical. If the data were symmetrical then the mean and median would be equal. I think it's good to know that skewed means off center in reference to that being centered.

Opal did not have much experience with statistics or collegiate coursework, but she was able to articulate her visualization of how the mean and median location would change with respect to changes to a distribution. Perhaps even more important, she believed it was important to highlight differences between representations, stating that "just telling me this is skewed doesn't really mean anything."

Interviewees were coded for a variety of responses to how they would teach the concepts needed to complete the Fuel Task to help prepare students for the AP-Statistics exam (see Table 4.29). Similar to the Tips Task, five interviewees stated they would just use the same pedagogical tactics they had described in the previous questions about how to teach a class of 20 students. Two respondents stated they did not know how to teach the content in the Fuel Task to prepare students for an AP-Statistics exam. Three major categories of suggestions emerged: long-term suggestions, content focused suggestions, and specific AP statements. These three categories were compared to KCS statement codings of traditional pedagogical, reform pedagogical and transnumeration teaching techniques to analyze trends between knowledge categories.

Table 4.29

Comparison across KCT Suggestions to Prepare Students for the AP-Statistics exam

	Traditional Pedagogical	Reform Pedagogical	Transnumeration
-I don't know	Jack, Amy		
Long-term Suggestions			
-Revisit Content			*Cole*
-Break-down content further	Lisa, Will, *Chas, Julie*		
-Teach to be analytical	*Hope*	*Hope*	
Content Focused			
-Know graph markings	*Kira*		Ben
-Know equations	Kadi, *Katilin, Kira*	Sally,	
-Highlight differences between graphs	*Julie*	Mia, Kent	Opal, *Jay, Kai*
-Relate skewness and median/mean		Neil, Myra	Neil, Carl, Sean
-Use applications	Will	Eve, Avis	
-Do more topic specific problems	Mary, *Dean*	Eve, Nick, James	Carl, Nick, Suzy, *Ava*
Specific AP Statements			
-Be an AP grader		Eve	
-Use multi-part AP questions	Kadi		

Note. An *italicized name* indicates the student answered the Fuel Task incorrectly, a regular font name indicates the student answered the Fuel Task correctly.

Long-term suggestions focused on the general knowledge needed to successfully answer questions about the Fuel Task's content. Of the six interviewees who provided a long-term suggestion, only two responded correctly to the Fuel Task. Five of the six interviewees were coded as suggesting a traditional pedagogical technique for teaching a class of twenty students, which shows a general discomfort in answering this interview question in an advanced manner.

Content focused suggestions ranged across a variety of topics: know specific graph markings, know equations, highlight differences between graphs, relate skewness to a mean or median, use applications *and* do additional problems. The content focused category was the most common coded response for interviewees, with 23 respondents suggesting at least one of these subcategories. Within this content subcategory, only three of the six respondents who made statements coded for a traditional pedagogical teaching tactic answered the Fuel Task correctly. In comparison, all nine respondents who made statements coded for a reform pedagogy teaching tactic answer correctly and respondents coded for statements in the transnumeration category answered correctly seven out of ten (70%) times. One plausible reason for this trend was as the series of eight questions on the interview protocol was asked, respondents who were weak in CKC were still focused on learning the content themselves and could not articulate meaningful responses *with teaching the content* to KCS and KCT questions.

The last category of suggestions by respondents with respect to the Fuel Task were AP-Statistics techniques used by teachers in the classroom. Two suggestions were made: to be an AP grader and to prep students for multiple-part questions with this material. Eve was a Beta that did well on the survey, getting four tasks completely correct and six at least partially correct. Eve talked about her high school experiences with AP-teachers:

> I know some teachers in my high school were AP graders and they knew the kinds of questions that were going to be asked on the test like the ones that are always asked and the ones that are like rarely asked. So I know boxplots, describing the skewness, describing the center, that is always a question on the AP-Statistics exam. I wouldn't focus on things that may not be on it. If you focus things that are going to be on it would help them succeed.

Eve was able to describe not only the general idea of teaching-to-the-test, but she brought in specific statistical terms that were focused on in her experiences. Kadi's answer to the Fuel Task was completely correct, but interestingly was coded for suggesting traditional pedagogical tactics. One reason might be because this was her first year in college and she had limited exposure to pedagogical tactics. When talking about how she would teach this material to prepare students for the AP-Statistics exam, Kadi said:

> I think there are a lot of different stuff that an AP-test could ask (with this graphical representation) because there are three types of graphs and three different ways you can get the information. They need to know a lot of the math behind it, like a lot of different formulas because I feel like they would ask five different questions on reading graphs. So I feel like knowing the formulas would help.

Kadi focused her response between thinking about the content (showing CKC) and thinking about how to teach the content (showing KCT) to help students be successful on the AP-Statistics exam. Statements that combined knowledge categories were rarely made by respondents. Most interviewees talked about content or pedagogy, but not both at the same time. Kadi's ability to discuss both content and pedagogy probably helped her become one of only two respondents to get the Fuel Task completely correct.

Fuel Task belief responses.

The preservice teachers beliefs about their ability to complete the Fuel Task were comparatively less positive than the other tasks. With the Fuel Task having two of the five answer choices correct, respondents had a good chance of reasoning to at least a partially correct answer. Kent was an example of a respondent who stated he used process of elimination to get to an answer, but arriving at the partially correct answer of "B." Interview data was coded to highlight respondent's limitations in different content areas that influenced Fuel Task responses. Figure 4.10 shows the comparisons of the mean response of all belief questions compared to the Fuel Task. Respondents viewed the Fuel Task as a more difficult task than the average task in the survey.

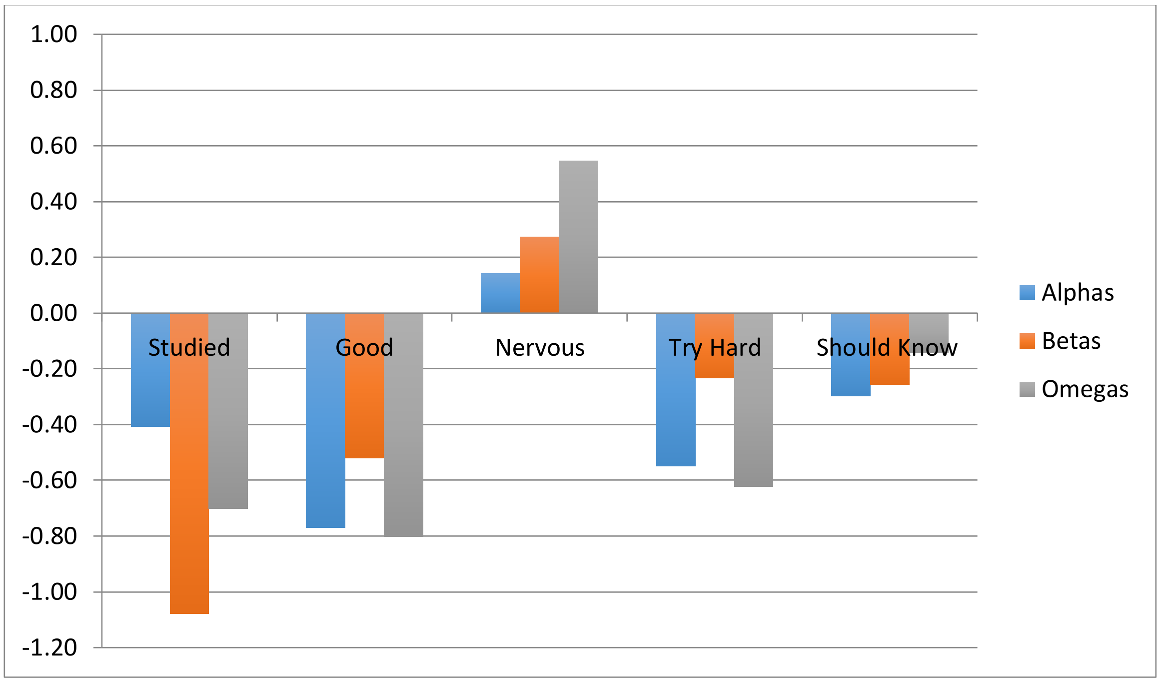


Figure 4.10. Comparisons of the mean belief rankings for all tasks to the mean ranking for the Fuel Task. Belief questions labeled above are as follows: Studied: This question is about a topic I have studied in a college class; Good: I am good at answering questions like this one; Nervous: I often feel nervous when I try to answer questions like this one; Try Hard: If I try hard, I can usually figure out questions like this one; Secondary: Secondary mathematics teachers should know how to answer this question.

In terms of opportunity to study this content in a college course, Beta's ranked the Fuel Task more than a point lower than they ranked other tasks and Omega's ranked the content lower by 0.70. Respondents also felt like they were less good at this task, they were more nervous than on other tasks, and they did not think if they tried hard they would answer this task correctly as much as other tasks. However, even though many respondents did not believe they had the skills to do this task, respondents did, overall, believe this material was important to know. When asked whether secondary teachers should know this material, responses averaged closer to the rest of the tasks (-.30, -.26, -.14), suggesting that even though respondents did not believe they had the knowledge needed in the Fuel Task right now, they should by the time they start teaching.

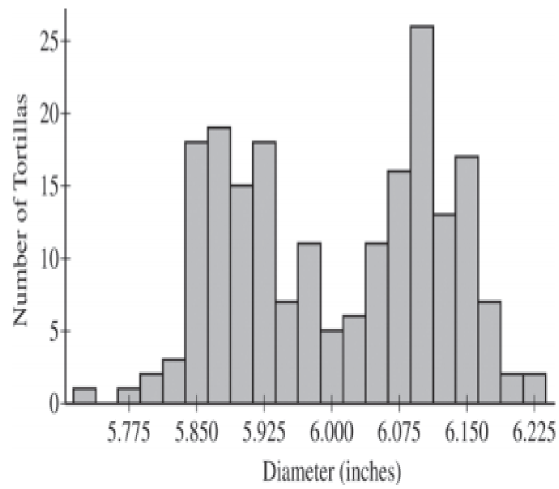
Factory Task

The Factory Task was chosen for the survey because all four expert reviewers indicated that transnumeration was a key to a reasonable solution. Additionally, the task averaged a score of 2.75/3.0 (between "a great deal" and "quite a bit") to the question about whether interview discussions about the Factory Task would involve transnumeration. The main knowledge categories the Factory Task was believed to target were SKC and KCS, providing research about two different knowledge categories. The only research question that the majority of reviewers felt the Factory Task would be of questionable value was 2a. "How do preservice teachers suggest they will use transnumeration as an inservice teacher?"

The Factory Task's graphical representation was a bimodal histogram (see Figure 4.11). The context of the question described a factory with *multiple production lines* of tortillas with a mean that was very close to the desired diameter of 6 inches, highlighting the importance of understanding the context behind the histogram. The Factory Task was challenging because the distribution was bimodal so the production lines most likely were not producing tortillas close to the mean, but rather the smaller and bigger errors were cancelling each other out. Answer "D" specifically required respondents to use transnumeration in thinking about the characteristics of a normal and bimodal distribution, possibly in a visual-overlapping picture. Transnumeration for the Factory Task could also occur when respondents think about where a mean-line was on the distribution. With two distribution types and hidden mean-lines, the Factory Task presented different opportunities for respondents to transnumerate the given graphical representation.

Peggy is a high school student in your class working on the following problem:

A factory's goal is to produce tortillas at 6 inches in diameter. The quality control workers in the factory take a random sample of 200 tortillas across multiple production lines. Results show a sample mean of 6.02 with the following distribution. Peggy describes the factory's performance saying "The factory is doing a great job of meeting their goal. The sample mean is only .02 inches away from the population mean goal." What should you do next as Peggy's teacher to further her understanding of the factor's performance?



A. The factory's performance is outstanding being only .02 inches away from the mean. As her teacher, I would agree with Peggy and ask if she thinks the company should take another random sample of 200 tortillas to verify results.
B. As Peggy's teacher, I would try to highlight why a histogram is used to represent data. Have Peggy create a Stem and Leaf plot by hand to see how the histogram above was developed.
C. Calculate the mean and median of the data set, and then ask Peggy to compare how the two measures relate to each other. Finish by highlighting the critical differences between the mean and median.
D. Have Peggy draw a normally distributed histogram and says it's mean is also only .02 inches away from 6 inches. Then ask which production line in the factory has better performance based on the two distributions.
E. Peggy should change the scale on the y-axis from an interval of 5 to 1 to get a more precise histogram of the amount of tortillas with each diameter range. Then ask Peggy to explain if the sample mean changed at all given the more detailed histogram.

Figure 4.11. The Factory Task and answer choices, answer "D" is the correct response.

The Factory Task positioned respondents to think like a teacher through a student comment about the factory's performance and specific pedagogical suggestions in answer choices. Answer "A" was designed as a distractor for respondents whose statistical experiences focused on measures of central tendency rather than other important aspects of distributions (i.e., spread). Answer choice "B" related histograms with stem-and-leaf plots, which actual align nicely in displaying datasets. Stem and leaf plots use exact numbers, but they preserve the shape of a distribution pre-grouped in base ten. Answer "B" was an incorrect pedagogical response because creating a stem-and-leaf plot with 200 tortillas would take the student way too much class time. Answer "C" was a distractor about measures of central tendency (similar to "A"), but not acknowledging that Peggy had translated information on the factory task correctly. Answer choice "D" was the correct answer for the Factory Task. Comparing the normal and bimodal distributions, Peggy should see that more tortillas were closer to the mean of six because the larger histogram bars occur in the middle of the distribution. Visually, Peggy should be able to see the area difference in the histogram bars and realize bars that were in the center should be close to six if the factory was in fact doing well. Answer choice "E" focused on the importance of scaling of the graphical representation rather than understanding the problems with the factories performance. Table 4.30 shows the responses for the Factory task sorted by the experience groups.

Table 4.30

Results by group for Factory Task

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)	Totals	Percentage
A	1	1	1	3	8.1%
B	4	2	1	7	18.9%
C	5	3	4	12	32.4%
D[a]	3	3	5	11	29.7%
E	0	0	0	0	0%
A, C & D	0	1	0	1	2.7%
B & C	0	1	0	1	2.7%
C & E	0	0	1	1	2.7%
B, D, & E	0	0	1	1	2.7%

Note. [a] denotes the best (i.e., completely correct) answer.

Answer “A” was selected 8% of the time as the only answer, and 11% of the time overall. Answer “A” was the fourth most popular choice, suggesting that respondents knew Peggy’s statement was incorrect and the factory was in fact doing poorly. Answer “B” was the second most commonly picked distractor with 19% responding to “B” only and 24% picking “B” as at least part of the answer. Alphas chose answer “B” most often between the different groups (57%) suggesting inexperienced respondents leaned towards stem-and-leaf plots as being important to highlight as a teacher in this situation.

Answer “C” was the most common response overall even though it was a distractor, with 32% answering only “C” and 40.5% answering “C” as at least part of their answer. Respondents were drawn to focusing answers to the measures of central tendency perhaps because of their familiarity with mean and median. Answer choice “C” uses the word *mean* three times in the question making it an attractive answer choice for students putting an overemphasis on measures of central tendency relative to spread. Answer “E” was never chosen as the sole answer of the task, and only answered as part of the solution twice (both by Omegas). When discussing answer choice “E” in interviews, respondents thought this teaching action was unrelated to solving the problem at-hand. Respondents who answered it as part of the answer believed changing the scale would make details of the graphical representation more explicit.

The second most popular response was the correct answer, letter “D.” Respondents chose “D” alone 30% of the time and as at least one of the correct answers 35% of responses, with the most respondents coming from the Omega group. The following section discusses the twelve correct answer justifications to provide additional insight into what respondents were thinking when they choose answer “D.”

Correct respondents' justifications for the Factory Task.

Alpha respondents struggled with the Factory task with only three of the 13 (23%) respondents answering "D." Alpha respondents chose distractors "B" and "C" more often than the correct answer. Beliefs about the Factory Task from correct Alpha respondents reaffirmed a limited ability to understand how (if at all) a bimodal distribution and measures of central tendency relate. When asked whether they agreed with the statement "I am good at answering questions like this one," the three correct Alpha respondents averaged a 3.3 response out of 7, close to the "somewhat disagree" score of three (see Table 4.31).

Table 4.31

Correct Alpha Student Justification for Factory Task

Respondent	Good[a]	Justifications
Jill	2	Jill did not interview but her reflection said answer "D" related the most to the factory's performance.
Suzy	3	Suzy knew they should be getting a normal distribution back, and comparing the bimodal with a normal distribution would highlight the factory's downfalls.
Cole	5	Cole knew the mean was not showing the factory's most common outcomes.

Note. [a] Good column is the response to statement, "I am good at answering questions like this one" where 1 is strongly disagree and 7 is strongly agree.

Only four of the 11 (36%) Beta respondents answered the Factory Task at least partially correct (see Table 4.32). This was a small improvement over the Alpha groups' average of 23% correct. Beta respondents who answered correctly averaged just above the neutral response at 4.25/7 to the belief question about whether or not they are good at answering this task. One of the four correct Beta students, James, actually answered "A", "C", and "D." During his interview, James justified answer "A" as a way to "look a little bit deeper" into the dataset and confirm or disaffirm Peggy's statement. James answered "C" because he thought it would highlight differences between mean and median in the dataset, which in his words was "a little all over the place." James stated for the Factory Task that having part of answer correct was reason enough to chose the answer choice.

Table 4.32

Correct Beta Student Justification for Factory Task

Respondent	Answer Choice	Good[a]	Justification
Kira	D	6	Kira knew the factory's performance was not good because the spread of the distribution was bimodal
Eve	Originally "B" changed to "D"	4	Eve realized through the interview that answer "D" would help Peggy see differences between a bimodal and normal distribution.
James	A, C, & D	5	James answered A, C, and D because he saw part of each answer as being important for Peggy's understanding.
Kadi	D	2	Kadi noticed there were multiple production lines during the interview and knew data should cluster near six.

Note. [a] Good column is the response to statement, "I am good at answering questions like this one" where 1 is strongly disagree and 7 is strongly agree.

Six of the 13 (46%) Omega preservice teachers were at least partially correct (Table 4.33). Correct Omega respondents were much more confident than Alpha and Beta students, averaging 5.17/7, just above agree to the statement "I am good at answering questions like this one." Two of the correct Omega respondents actually changed answers from "C" to "D" during interviews. Recall answer choice "C" was the most common final answer, and would have been even more common if respondents like Kai and Jeremy had not changed their answers during interviews.

Table 4.33

Correct Omega Student Justification for Factory Task

Respondent	Answer Choice(s)	Good[a]	Justification
Carl	B, D, & E	7	Carl wanted Peggy to focus on the measurement unit (grams) in the task
Nick	Changed A & D to only D	5	Nick saw correct aspects from both A and D and combined them into his teaching suggestion for Peggy.
Kent	D	3	Kent thought comparing two distributions would help Peggy see a better distribution.
Kai	Changed "C" to "D"	7	Kai believed knowing a mean and median was not enough information to evaluate if the factory was doing well.
Jeremy	Changed "C" to "D"	5	Jeremy knew that if you cut the distribution in half, the distribution had a normal shape.
Sean	D	4	Sean realized the factory was doing poorly even though the mean looked good.

Note. [a] Good column is the response to statement, "I am good at answering questions like this one" where 1 is strongly disagree and 7 is strongly agree.

Interview discussions promoting change.

As previously noted, interview discussions provided opportunities for preservice teachers to improve knowledge. For the Factory Task, four respondents (Eve, Jeremy, Nick, and Kent) changed to the correct answer of "D" during interviews. Eve and Jeremy were paired together and both arrived at answer "D" through their answer discussion:

> Eve- I do think it is important to know the difference between the mean and median because a lot of people can get that confused but in this case I don't know if that is going to…
>
> Jeremy- Now that looking at this I don't know if C would be helpful.

> Eve- D, D is it?
>
> Jeremy- Yes, D.
>
> I - Okay, why D?
>
> Eve- Because answer D has her seeing the difference between the normally distributed histogram and whatever this distribution is (bimodal).

Kent talked about how he picked answer "C" because a mean could be close to the middle of the distribution if the quantities on each side of a mean balance each other out:

> I was looking at how they want to find the average is close to six, but that does not necessarily mean that any of the tortillas are necessarily six inches in diameter. Because when you could have 100 at eight inches and 100 at four inches the average is going to be six, but none of those are six (inches). I wanted to know the difference between the mean and the median to show that's not necessarily true.

Kent changed his answer from "C" to "D" after Kai mentioned how a normal distribution should have more data points in the middle of the display near the mean and median. Kent understood the idea that the bimodal distribution was symmetrically balanced, but initially struggled to focus on how this task was really focused on how one views the spread of a distribution independent of measures of central tendency.

Content knowledge about graphical representations.

Interpreting how measures of central tendency appeared on a bimodal distribution proved to be challenging as many respondents focused on the distribution shape or center as their main justification for answer choice. It could be that responses were influenced by the uniquely different answer choices of the Factory Task. The three most common responses involved very different benchmark ideas in statistics, with answer choice "B" focusing on an alternative way to display (i.e., stem and leaf plots), answer "C" focusing on popular measures of central tendency, and answer “D” relying on respondents’ ability to transnumerate. Respondents were also coded for making statements about not knowing the answer, the performance of the factory being good or bad, and specific contextual elements in the task. Results are shared below.

Center focused statements.

Respondents who were coded for making a statement about measures of central tendency for the Factory Task’s distribution (mean or median) did very poorly on the task overall (see Table 4.34) with only two of the fifteen (13%) respondents getting any part of the task correct. Eve was the only respondent to answer the task correctly while making this statement, and her coded statements about calculating the mean and median occurred before she changed her answer from “C” to “D.” Similarly, only one of the five respondents who stated the mode was the tallest bar actually answered the Factory Task correctly.

Table 4.34

Center-focused Statements

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
Center Statements			
-Calculate mean & median	*Opal, Jay, Neil, Jill, Avis, Lisa, Mia*	*Sally, Jack,* James[a], *Ben,* Eve	*Mary, Kelly, Hope*
-Mode is tallest bar	*Mia, Will*	*Ben, Ava,* James[a]	
Center to Spread Statements			
-Correct statement	Suzy	Eve	*Myra, Hope,* Jeremy
-Incorrect Statement		*Ben*	
-Mean should be a taller bar	*Dean,* Cole,	*Ava,* James[a], Kadi, Taft	*Julie, Hope, Myra,* Kai

Note. [a] represents a partially correct answer. An *italicized name* indicates the student answered the Fuel Task incorrectly, a regular font name indicates the student answered the Factory Task correctly.

There were a variety of statements made by respondents that attempted to use some sort of knowledge of the mean or median to make a statement about what was happening with the spread of the distribution. Ben was a Beta respondent who answered the Factory Task incorrectly and was trying to make links between measures of central tendency and the distribution. Ben said, “I think Peggy should do this because if the mean and median are close it proves there are no outliers altering the data.” Ben’s statement is true for some datasets, but there are also cases where outliers could exist in a distribution while the mean and median are the same (e.g., Figure 4.12).

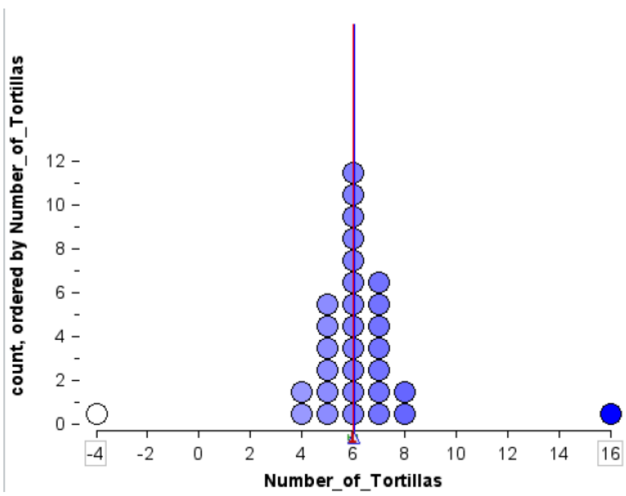


Figure 4.12. An example of a distribution with data points that are likely to be outliers, this figure shows two data entry points entered incorrectly (-4 and 16), while mean and median lines for the dataset are exactly the same (seen as blue and red lines at 6 inches).

A common statement during Factory Task justifications was that the mean should be a taller histogram bar compared to the rest of the histogram bars. Ten respondents were coded for this kind of statement, with four (40%) responding at least partially correct. Some respondents were able to make accurate statements about the relationship of measures of central tendency and a distribution. Suzy was an Alpha student whose statement was coded as a correct statement about a measure of central tendency towards the spread. Suzy highlights how you can find a center (median) on a distribution:

> Suzy – Um, you could figure out the median if you really needed to know that.
>
> I – Ok, how would you do that?
>
> Suzy – Um, theoretically because you know how many are in each one of these bars, I don't know if there is an easier way to do it like in elementary school when you wrote them (numbers) out and crossed off each end, theoretically I guess you could do that.

There were other interesting statements that used center to infer the spread. Myra was an Omega student who concluded that the mean and the median would be about the same because the bimodal distribution was symmetrical. Eve and Jeremy realized that where the median was in relation to the mean could show whether the distribution was skewed right or left. Overall when respondents attempted to relate a measure of central tendency to the spread of the distribution, it provided an opportunity to discuss the inner-workings of the task.

Spread focused statements.

Some respondents made statements that focused on the spread of the distribution explicitly while discussing the Factory task (see Table 4.35). Two respondents stated that they did not know what a normal distribution was even though both of the respondents had completed their collegiate statistics course. Eight students talked about how a stem-and-leaf plot helps someone visualize a distribution and believed this was very important, although only one (13%) actually answered the Factory task correctly.

Table 4.35

Made Spread Focused Statements for the Factory Task

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
-I don't know what a normal distribution is.		*Chas, Amy*	
-Stem & leaf help visualize	*Klara, Bell, Dean*	*Sally, Chas,* Eve	*Julie, Mary*
Spread to Center Statements			
-Many data points are far from mean	*Will, Mia*	*Jack,* Kadi	Kent,
-Without symmetric distribution, mean is difficult to use		*Ava*	Carl[a], Kent

Note: An *italicized name* indicates the student answered the Factory Task incorrectly, a regular font name indicates the student answered the Factory Task correctly. [a] denotes a partially correct response.

Ava was very close to answering the Factory task correctly, perhaps because she had taken AP-Statistics in high school, but did not remember all the necessary content to answer correctly. Ava knew the mean was not a great measure to use as a predictor with a distribution that does not approach normal, but she did not make any statements about the Factory's performance across the histogram. Other respondents (Carl and Kent) made statements about not drawing inferences with a mean in a non-symmetric distribution.

Four common trends involving the context of the task emerged while discussing results of the Factory Task that did not show explicit language of the relationship between center and spread (see Table 4.36). These four trends were realizing there was a total sample size of 200 tortillas, that there were multiple production lines, that resampling was a good idea and that the distribution should be normally distributed. The first category had seven respondents who realized there was a 200-tortilla sample, of which only three (43%) answered correctly. Two of the five respondents who suggested resampling answered the task at least partially correct.

Interestingly, no Alphas suggested resampling perhaps because of their limited exposure to statistics.

Table 4.36

Additional Trends in the Factory Task

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
-200 tortillas	Cole, *Neil, Lisa*	Kira, *Amy*	*Mya,* Kai
-Multiple production lines		Kadi	Nick,
-Resample		*Sally, Ava,* James [a]	*Mya,* Nick
-Should be symmetric	*Opal, Avis,* Suzy, Cole	Kadi, Taft, Kira	*Kelly, Chloe,* Nick

Note: An *italicized name* indicates the student answered the Factory Task incorrectly, a regular font name indicates the student answered the Factory Task correctly. [a] denotes a partially correct response.

When respondents stated the distribution should be symmetrical, four out of ten (40%) actually answered the task incorrectly. Given that answer "D" was the only answer that used the *normal distribution* terminology, it suggested that perhaps respondents' conception of a normal distribution was limited to a general picture of a bell-curve, or was lacking depth in investigating differences between distribution attributes like center and spread. Perhaps respondents viewed symmetric distributions as typical and did not know how to search the dataset for why it did not look symmetrical. There appeared to be a gap between respondents being able to state what a distribution should look like and being able to reason why a distribution may not embody that shape.

Transnumeration in the Factory Task.

Perhaps the most important responses during interviews for the Factory Task were those that used transnumeration to justify answer choices. One category that described how respondents solved the Factory Task was the transnumeration in context category (see Table 4.37). Two subcategories of statements were made about the transnumeration in context category: statements about the bimodal spread showing poor performance for factory and statements about how the mean and median show the factory performance doing well, but the spread of the distribution shows otherwise. Twelve respondents were coded for statements using transnumeration in context, with 11 (92%) at least answering partially correct. Chloe was the only respondent who answered incorrectly. In her reflection section, she wrote: “Although the sample mean is only .02 away from its goal, the histogram illustrates the company has some work to do. A bell curve would be nicer, more tortillas at the actual goal.” Unfortunately, Chloe did not participate in an interview as it would have been interest probe her reasoning in choosing answer “C” given she realized the distribution was not ideal for the factory’s performance.

Table 4.37

Transnumeration in the Factory Task

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
Transnumeration in context			
-Bimodal implies poor performance	Cole	Eve, Kira	*Chloe*, Carl[a], Sean, Kai, Kent
-Center shows good performance, spread shows otherwise	Cole	James[a], Taft	Nick, Sean
Transnumeration between Graphs			
-Comparing bimodal and normal distribution is helpful	Suzy, Cole	Kadi, James[a]	*Kelly*, Carl, Nick, Kai, *Myra*
-Comparing different types of graphs is helpful	*Opal*		*Julie*, *Mary*

Note: An *italicized name* indicates the student answered the Factory Task incorrectly, a regular font name indicates the student answered the Factory Task correctly. [a] denotes a partially correct response.

Respondents who knew the bimodal spread was bad for factory performance were able to apply knowledge about the distribution to make a judgment about factory performance. This was typically done when respondents answered correctly, with seven out of eight (88%) getting the task correct. The other transnumeration within the context subcategory involved statements about how the measures of central tendency depicted the tortilla production going well, but the spread of the tortilla distribution showed poor production performance. All five respondents coded for statements in this subcategory answered the Factory Task at least partially correct. Keeping track of the context, spread and center was one of the most complex thought-processes preservice teachers needed to do while completing the Factory task. However, a lack of using context to drive statements about spread or center resulted in misunderstandings and incorrect responses.

Transnumeration between graphs was the other category with two subcategories (i) comparing bimodal and normal distributions is helpful *and* (ii) comparing different types of graphs is helpful (i.e. stem-and-leaf plot versus histogram). Respondents who saw value in comparing normal and bimodal distributions answered correctly seven out of nine times (78%) including two of the three Alphas who answered the Factory Task correctly. The two Alphas, Cole and Suzy, both saw value in comparing graphs when teaching and learning. This direction of response was perhaps an entry point for more advanced statements about comparing center, spread, and the context of the Factory task. Cole showed some advanced knowledge in his interview:

> Cole – I guess something I noticed was that the application of the problem the goal is accuracy to diameter of 6 inches, and looking at the graph it does not look like it is super accurate, but the mean and likely the median show that the numbers alone might be. So I thought it would be interesting to compare with a normally distributed histogram.
>
> I – Why would you want to do that comparison, or why would you do that?

> Cole – To show that we could have two of the same mean and one would be far more accurate than the other and make that distinction. Show how an accurate mean does not necessarily mean overall accuracy in terms of the application.

All three respondents who made statements coded about there being value in comparing different types of graphs like a stem-and-leaf plot and a histogram answered correctly. With nearly a third (12 out of 37) of respondents coded for transnumeration between graphical representation statements, respondents believed making comparisons between graphs would help facilitate student learning even if they had limited content knowledge.

Markings that supported transnumerative thinking.

Markings on the Factory Task were less common than on the Fuel Task, although respondents were active in underlining information and crossing off answers they believed to be incorrect (see Table 4.38). Eleven respondents made at least one kind of marking. Only five of these respondents (45%) answered the Factory Task at least partially correct, which was only slightly above the 35.1% of respondents that answered the Factory Task at least partially correctly overall. This suggested that underlining, crossing off incorrect answers, and marking on the distribution did not really help respondents find the correct answer.

Table 4.38

Markings in the Factory Task

	Alphas (n=13)	Betas (n=11)	Omegas (n=13)
-Underlining text	*Opal, Will,* Jill [a]		*Kelly,* Nick
-Crossing off wrong answers	Jill [a]	*Ava*	*Mya, Chloe* [a], *Kelly,* Nick, Sean
-Marking on distribution		James [b], Kira	*Mya,* Nick

Note: An *italicized name* indicates the student answered the Factory Task incorrectly, a regular font name indicates the student answered the Factory Task correctly. [a] denotes a partially correct response. [b] denotes a partially correct response.

There were instances where marking on the distribution showed content knowledge beyond merely completing the task. James was one of only four students to mark on the histogram. James circled the low histogram bars near the center of the distribution (see Figure 4.13 below), suggesting as he discussed in his interview that the center of the distribution was showing poor factory performance because there should be a large number of tortilla values positioned here. Recall James answered the task partially correct, but was able to make some advanced statements about transnumeration and how the factory should resample to see if they got a similar distribution.

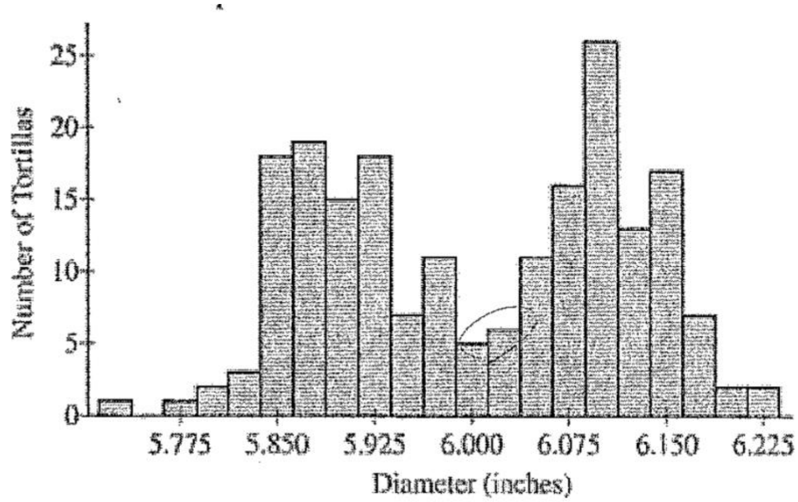


Figure 4.13. James's markings of histogram lengths while taking survey.

Mya was an Omega who only completed the survey. She marked a line in the middle histogram bar (see Figure 4.14 below) and wrote in her reflection section,

> I chose A because while .02 does not seem to be much, looking at the histogram it is clear that a lot of tortillas exceed the ideal 6 inch diameter requirement. By looking at 200 more tortillas, I think a better conclusion can be reached about the factory's performance.

Mya marking suggested she used the line as a kind of threshold of acceptable values for the factory's performance.

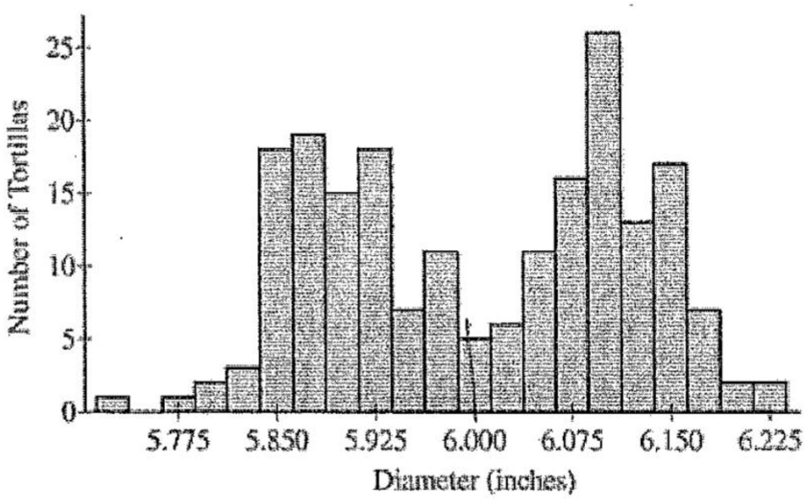


Figure 4.14. Mya's marking the middle histogram bar on the Factory Task.

Kira and Nick were the only students who drew a bell curve shape on their survey. Kira actually drew a bell curve in her reflection section (see Figure 4.15).

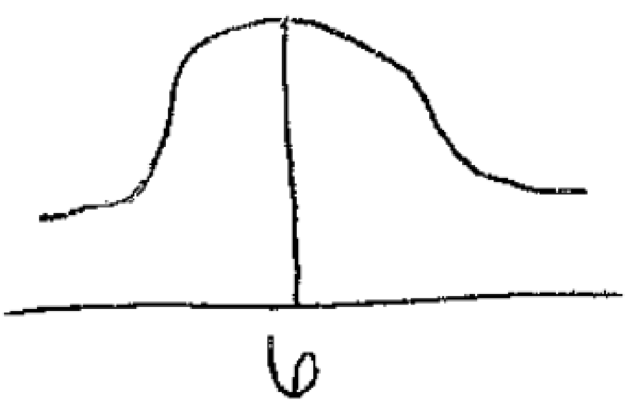
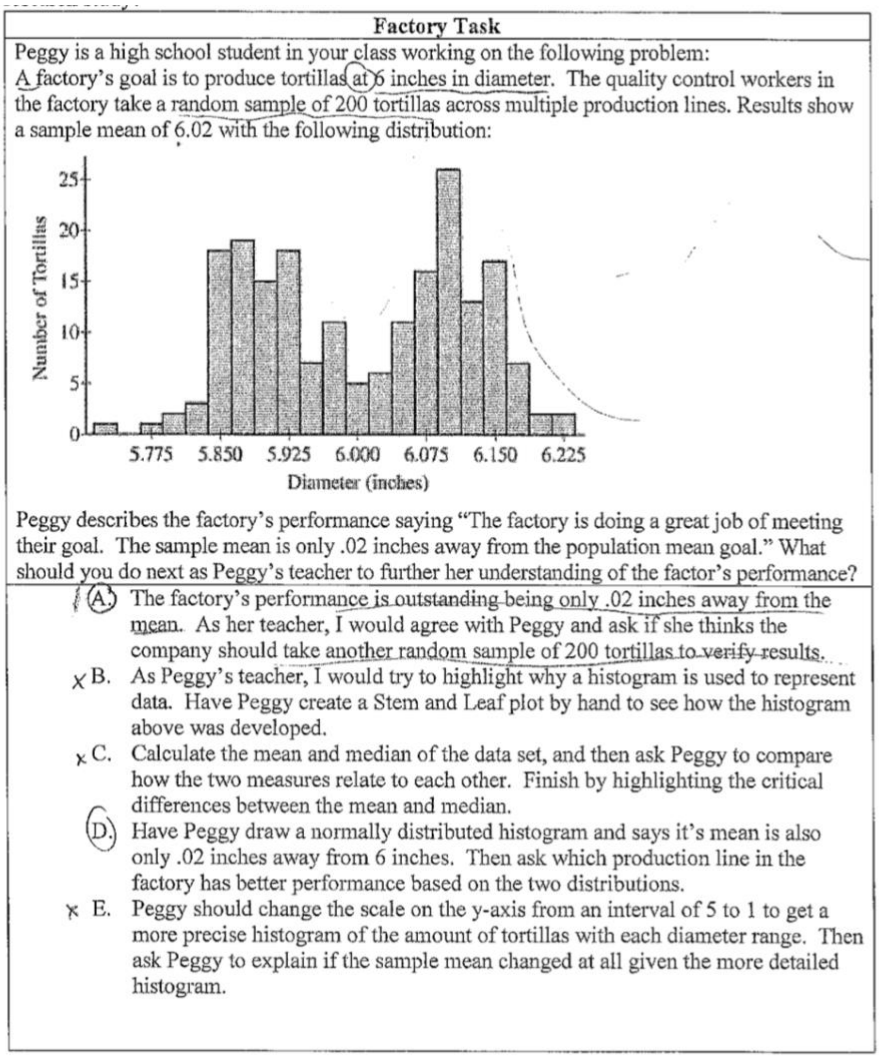


Figure 4.15 Kira's marking in the reflection section of the Factory task.

Kira was able to not only show the shape of expected distribution, but also drew a mean line on the distribution. During here interview, Kira made some strong statistical statements. She justified her correct answer by saying:

> I answered "D" because you would want a normal distribution to see most of your products are right where you want them at the 6, and this is not a normally distributed graph that they have here. And so when it says Peggy "The factory is doing a great job of meeting their goal." I was like no I don't think they are because they have different measurements all the way around. So most of them would want to be 6, that is why I said that.

Nick was one of the most knowledgeable respondents who took the survey in that he tied for most tasks at least partially correct with seven and also tied for the most tasks completely correct with five. Nick's markings on the Factory Task were more active than any other respondent (see Figure 4.16). Nick circled and underlined words during the task and throughout answer choice "A." He crossed out B, C, and E when he eliminated them. Probably the most interesting marking Nick had was tracing a faint bell curve over the right side of the distribution along with tracing a bell curve in general off to the right-hand side of the figure. Nick's tracing during the survey showed his comprehension of how the Factory Task was concerned with understand the spread of tortillas.



Factory Task

Peggy is a high school student in your class working on the following problem:
A factory's goal is to produce tortillas at 6 inches in diameter. The quality control workers in the factory take a random sample of 200 tortillas across multiple production lines. Results show a sample mean of 6.02 with the following distribution:

Peggy describes the factory's performance saying "The factory is doing a great job of meeting their goal. The sample mean is only .02 inches away from the population mean goal." What should you do next as Peggy's teacher to further her understanding of the factor's performance?

A. The factory's performance is outstanding being only .02 inches away from the mean. As her teacher, I would agree with Peggy and ask if she thinks the company should take another random sample of 200 tortillas to verify results.

B. As Peggy's teacher, I would try to highlight why a histogram is used to represent data. Have Peggy create a Stem and Leaf plot by hand to see how the histogram above was developed.

C. Calculate the mean and median of the data set, and then ask Peggy to compare how the two measures relate to each other. Finish by highlighting the critical differences between the mean and median.

D. Have Peggy draw a normally distributed histogram and says it's mean is also only .02 inches away from 6 inches. Then ask which production line in the factory has better performance based on the two distributions.

E. Peggy should change the scale on the y-axis from an interval of 5 to 1 to get a more precise histogram of the amount of tortillas with each diameter range. Then ask Peggy to explain if the sample mean changed at all given the more detailed histogram.

Figure 4.16. Nick's markings on the Factory Task.

Interestingly Nick did originally answer "A" and "D" and later changed his answer to only include "D" because he disagreed with Peggy's statement. Originally Nick thought answer "A" was correct because it would lead Peggy to either a more normal distribution, or confirm the bimodal shape:

> Neither of them (i.e. answers "A" and "D") is perfect, but together I think they would be the best. So "A" is about the mean, which winds up being about 6.02, and asking her to take another sample. This sample should have been normal, is what I decided. Taking that many tortillas from a population should have given a normal result, or like a normal distribution. The fact that it was not means that it's the extreme rare case where it does

> not happen the information is not normal. So taking another sample would verify that its normal and she would get roughly the same mean, then she has the right answer, or it would suddenly everything would go wrong and she would get a very different answer because the data is not normalized. So "A" would allow her to see just because this data is close, the fact that its not a normal distribution means they probably should have different values for the mean and median.

Although answer choice "A" clearly has an incorrect statement (i.e., Peggy is right), Nick developed some strong reasoning in attempting to combine these two answer choices to help Peggy's knowledge develop. The next section presents results about respondents SKC during the Factory Task.

Specialized Knowledge of the Content in the Factory Task.

SKC was evaluated by asking respondents what knowledge teachers have with this material that other working professionals do not have. As was the case with the Tips and Fuel Tasks, it was difficult for interviewees to describe distinct differences between teachers and other professionals. Respondents mentioned different CKC commonly (see Table 4.39), similar to the Tips Task. Three categories of CKC coded statements emerged: generic statements, specific statistics statements, and transnumeration statements.

Table 4.39

Responses to SKC Questions that Focused on CKC

	Alphas	Betas	Omegas
Generic			
-Explain the "why"	*Will*[2], *Jay*, Cole		
-Read the graph	*Mia*	*Jack*[2], *Ben*[2], *Amy*[2]	*Hope*[2], Kent[2]
-Vocab or terms	*Avis, Lisa*	*Sally*[2]	*Mary*[2], *Julie*[2]
Specific Stats			
-Know spread	*Will*[2]	*Sally*[2]	
-Know normal distribution	*Klara, Dean*[3]	Eve[2]	Sean[3], Nick[3]
-Stem and Leaf Plot		*Amy*[2]	*Julie*[2]
-Mean and median	*Neil, Dean*[3]	*Jack*[2], *Ben*[2], Kira[3]	*Hope*[2], *Kelly, Myra*[2], Nick[3] Sean[3], Kent[2]
-Random sample		Kira[3]	*Mary*[2], Kai[2]
-Standard Deviation	*Dean*[3]		Sean[3], Nick[3]
Transnumeration			
-Compare mean and spread		Kira[3], Eve[2], James	Jeremy, *Myra*[2]
-Different sample & population means			Kai[2]

Note: Often preservice teachers made multiple subcategory comments, which are numbered for cross-table analysis. An *italicized name* indicates the student answered the Factory Task incorrectly, a regular font name indicates the student answered the Factory Task correctly.

The interview comments of eight individuals were coded for SKC (see Table 4.40), and broken into two categories: general and extra. The general subcategory focused on how to read and explain graphs to students and know standards by grade. Both Alphas and one of the two Betas whose comments fit into this category answered the task incorrectly. Chas was one of these four respondents that made a statement coded into the general category, but struggled to describe his knowledge with meaningful vocabulary:

Chas – Well even beyond that I guess, they need to know where they can go with this. Like aside from knowing the terms being able to see what it helps out the line in the future.

I – What do you mean by the line in the future?

Chas – Like further on, what can you learn from knowing this type of stuff.

I – Oh, ok, ok.

Chas – Have a further developed understanding.

I – Yes, I see what you're saying. Is there anything in particular that would be that future direction?

Chas – Mmm, no (laughing).

In contrast, all four respondents who were coded with statements in the extra category answered correctly. Two of these respondents were Betas and two were Omegas. More interestingly, all of these respondents were coded for CKC statements that were either specific statistics or transnumeration categories (see Table 4.39), suggesting that as CKC advanced, so did SKC.

Table 4.40

Statements Showing SKC for the Factory Task.

	Alphas	Betas	Omegas
SKC - General			
-Read and explain graphs	*Klara*	Kadi	
-Know standards by grade	*Opal*	*Chas*	
SKC - Extra			
-Read the graph deeper		Kira	Nick
-Lead students to reasoning		Eve[a]	Jeremy[a]

Note. An *italicized name* indicates the student answered the Factory Task incorrectly, a regular font name indicates the student answered the Factory Task correctly. [a] Shows a changed answered to correct during interview.

The next two sections describe the pedagogical knowledge coded from interviews about preservice teachers' KCS and KCT. KCS conversations typically were responses to a comparison question where respondents were asked to think about the difference between a high school student's knowledge of the content and their own knowledge. KCT results focused both on teaching a class of students and helping students prepare for an AP-Statistics exam.

Knowledge of content and students.

The Factory Task was challenging for respondents to show KCS particularly because of the very different knowledge was involved to justify answers B, C, and D, but answers were chosen pretty evenly at 24.3%, 40.5%, and 35.1% respectively. Interview results showed respondents who answered incorrectly often did not believe there would be a difference in how they, as undergraduate students, viewed the Factory Task compared to a high school student. Responses were coded into three categories: students dictate, no difference, and difference (see Table 4.41).

Table 4.41

Factory Task Response Differences Between Teacher and Student Knowledge

	Alphas	Betas	Omegas
Students dictate		*Jack*, Amy	
No Difference			
-General	*Klara, Jay, Neil, Avis, Lisa, Mia*	*Chas,*	Kai, Kent, *Mary, Hope*
-Histograms are simple	Cole *Opal, Klara*	Kira, *Ben*	
Difference			
-CKC or vocab	*Will*, Suzy	Eve[b], James[a]	*Kelly, Julie, Myra*
-Histograms are complicated	*Dean*	*Sally*	
-Context			Carl[a], Nick
-Transnumeration		Kadi	*Myra*, Jeremy[b], Sean

Note. An *italicized name* indicates the student answered the Factory Task incorrectly, a regular font name indicates the student answered the Factory Task correctly. [a] marks a partially correct answer. [b] notes a changed answered to correct during interview.

Of the 15 responses that were coded as no difference between the PST's knowledge and a high school student's knowledge, only four (27%) answered the Factory Task correctly. Subcategories for respondents who believed there would not be a difference in knowledge were in two areas: a very general statement about there not being a difference in knowledge and a statement about how the graph or histogram was not a complicated visual to decode.

Fourteen respondents stated there was a difference between their knowledge and a high school student's knowledge, or which eight (57%) answered the task at least partially correct. The difference category was broken into four subcategories: CKC or vocabulary, graph type, context, and a transnumeration difference. The most common difference pointed out by respondents was a difference about a specific statistical term (e.g., mean) on conceptual idea with seven of the respondents stating that would be different. In contrast to the respondents in the no difference subcategory already discussed, two respondents in the difference subcategory were coded for statements about how high school students would struggle to translate a histogram.

Perhaps the most advanced statements from respondents showing a development of KCS came from those who pointed out knowledge would be different between themselves and students in the context of the task and being able to transnumerate part of the task. Jeremy pointed out a difference in how he could transnumerate the distribution more than a high school student stating:

> I feel like a student that hasn't taken the stats class before would look at this as a normal graph. I think they would be closer to saying yes (i.e. Peggy is correct) because they haven't really had a stats class to know. I think a "normal" graph means to them a graph that could happen in real life. This doesn't look too out of the ordinary, you have a lot in the middle area up and down so that's like real life but I feel like a statistic student will know that it's not the normal bell-shaped curve.

Jeremy, although limited in his word choice to describe the situation, was able to talk about the distribution in the context of how it could happen in real-life. He linked the given histogram with the more likely normal distribution displaying his ability to transnumerate between the given bimodal distribution and the not-present normal distribution.

Knowledge of content and teaching.

Respondent's suggested different techniques to teach the concepts in the Factory Task to a class of 20 students, which were coded into three categories: traditional pedagogical, reform pedagogical and transnumeration (see Table 4.42). All statements made by Omegas were coded as reform pedagogical or transnumeration, suggesting that perhaps their coursework was changing their beliefs to think past traditional pedagogical methods. Traditional pedagogical suggestions focused on two areas: listing vocab or content and lecturing students. Only four respondents were coded for a traditional pedagogical style, and only one of these individuals (25%) answered the Factory Task correctly.

Table 4.42

Factory Task Statements that Demonstrated KCT

	Alphas	Betas	Omegas
Traditional Pedagogical			
-List vocab or content	Suzy, *Mia*	*Ava*,	
-Lecture	*Will*	*Ava*	
Reform Pedagogical			
-Group-work	*Klara*, *Jay*	Cole, *Chas*, Kadi, *Ben*, *James*	Kent, Kai, Nick, *Mary*, *Hope*, *Julie*
-Use questions techniques	*Klara*	Kira, *Chas*, *James*	*Hope*, *Kelly*
-Students make graphs	*Neil*	*Jack*	
-Real-life application	*Opal*, *Neil*, *Avis*	*Sally*	Jeremy[b]
Transnumeration			
-Compare histogram with another representation	*Lisa*	Eve[b], Amy	
-Give different histograms	*Dean*	Eve[b], Kadi	Carl[a], *Myra*

Note. An *italicized name* indicates the student answered the Factory Task incorrectly, a regular font name indicates the student answered the Factory Task correctly. [a] marks a partially correct answer. [b] notes a changed answered to correct during interview.

The majority of statements were coded into the reform pedagogical category for the Factory Task. The reform pedagogy category was broken into four subcategories: group-work, ask questions towards answers, students make graphs, and real-life application. The most common suggestion for teaching the content in the Factory Task was the use of group-work with students, with thirteen respondents suggesting this tactic. Twenty-one respondent statements were coded into a reform pedagogical subcategory, with seven (33%) answering correctly. Recall Kira demonstrated deep CKC for the Factory Task as she described, with correct vocabulary, differences between a normal distribution and bimodal distribution. Kira really focused only on discussing how she would use questioning to advance the content for the Factory Task when she responded to this question:

> I think definitely questions that lead them to what you want them to see. So first you would say "What do you notice about the graph?" Pretty basic and when you get his or her answer, when somebody says something that triggers the conversation, like if someone would say "This bar is super long." Then that would be your next leading question and you could say, "Well why do you think that?" and "How does that affect this graph?" Get them to the answer that the distribution is kind of skewed and that the company is not really making a lot of the 6-inch tortillas from their random sample.

Kira focused on trying to direct not dictate the conversation, which is a very important pedagogical tactic. However, Kira really did not have a complete view of what a lesson should look like with this task. Her comments showed she values reform teaching tactics, but has not experienced enough teaching to really describe a lesson thoroughly. Jeremy was an Omega respondent who was still developing his opinion of exactly how to teach this concept, but arrived at a hands-on tactic:

> I don't know exactly how to do it but I would try and make the numbers come to life. If possible, like make a cookie at each diameter and hand them out to the students and then you'd see some of these kids getting more than a 6-inch diameter and the other kids getting less. They will see how half of the diameters are failing. They would realize that the mean is at six but the diameters aren't all at six.

Of the five respondents whose statements were coded for using a real-life application to teach statistical concepts in the Factory Task, Jeremy was the only one to describe a specific situation tailored with student interests in mind (and the only one to answer the task correctly). Teaching while eating cookies would be a topic that students would really enjoy, but it also highlighted the context of the Factory Task well. The other four respondents believed students would benefit from real-life applications, but failed to describe a real-life application for a high school student.

The last category was the transnumeration category, which was broken into two subcategories: compare the histogram with another representation and give different histograms. Only four out of seven (57%) respondents with statements coded into the transnumeration category answered correctly. This was an interesting finding because some respondents who answered correctly could not articulate how to lead high school students towards the same answer. Some inservice teachers are described as having content knowledge, but not able to teach content. KCT findings here align with that statement. Respondents who understood there were distribution differences between what would be ideal (i.e., a bell curve) and what was present (i.e., the bimodal distribution) did not regularly respond with statements displaying strong KCT.

Interviewees provided different suggestions on how they would teach the statistical concepts in the Factory Task to help prepare students for the AP-Statistics exam (see Table 4.43). Five respondents stated they did not know how to teach this content to prepare students for an AP-Statistics exam. Five additional respondents made statements about not knowing how to teach in preparation for the AP-Statistics exam because they had never taken the exam themselves. Together then, 32% of respondents felt strongly that they were underprepared for teaching the statistics in the Factory Task! Other major categories emerged as suggestions to prepare for the AP-Statistics exam: general, content specific, and AP-specific statements. Statements were coded in these three categories and compared to statements coded in the KCS categories of traditional pedagogical, reform pedagogical and transnumeration (see Table 4.42).

This comparison weaves respondents' statements about teaching content and assessing content on the AP-Statistics exam.

Table 4.43

Comparing KCT Responses with Teaching Towards the AP-Statistics Exam Responses

	Traditional Pedagogical	Reform Pedagogical	Transnumeration
Unsure			
-I need more experience	Suzy	*Jay, Avis, Chas*	*Amy*
-I don't know	*Will*	*James, Mary, Kelly, Julie*	
General			
-Do examples		*Chas*, Kent	
-Practice vocab/ terms/ reading the problem	*Ava*	*Klara,* Kira	
-Break down the problem	*Mia*	*Ben*, Cole, Kai, Kadi	Kadi
-Teach critical thinking		*Hope*, Kira	
Content Specific			
-Teaching specifics about histograms	*Ava*	*Opal*, Kira, Nick	Carl[a]
-Know when to use each type of distribution		Jeremy, Klara	Myra, *Dean*
AP-Specific			
-Teach what's on the test	*Will*	*Neil, Jack*	
-Do previous exam questions		*Sally, Chas,* Jeremy[b]	*Lisa*, Eve[b]

Note An *italicized name* indicates the student answered the Factory Task incorrectly, a regular font name indicates the student answered the Factory Task correctly. [a] notes a partially correct answer. [b] notes a changed answered to correct during interview.

The general category included statements that involved little to no statistical vocabulary in descriptions. Respondents coded for statements in the general category struggled on the Factory Task, with five out of eleven (45%) answering correctly. Subcategories were to do this problem or similar problems in class, know vocabulary, terms, and how to read the problem, to break the problem down further than before, and to teach students to be analytical of critical thinkers overall. The subcategory about teaching students to be analytical or critical thinkers was interesting because this was a big-picture suggestion. Often big-picture suggestions are lost in high-stakes testing situations as teachers feel pressure to develop students for the now, a short-term gain, and not the long-term gain.

The second major category included content specific suggestions, where specifics to histogram or distribution aspects were a focal point of the response. Twelve students made statements coded in the general category, where only nine respondents made statements coded into the content specific category. Interestingly, six of the nine (67%) respondents making a content specific suggestion actually answered the task correctly, compared to four of the twelve (33%) answering correctly in the general category. This provides further evidence that aligned to the results from the Tips and Fuel Task that if preservice teachers can articulate specific statistical content about how they would teach material, they are much more likely to answer the task correctly.

The AP-specific category consisted of respondents' statements coded into two subcategories (i) teach what's on the test and (ii) do previous exam questions. Of the eight respondents who were coded for statements that fall into the AP-specific category, only two (25%) answered correctly and both changed to the correct answer during their interview. Respondents who provided this kind of statement were often struggling on the task's statistical content itself and went back to a secondary part of the question to focus on, specifically their ideas of what high-stakes tests entail.

When looking across respondents' statements coded throughout the three KCS categories of traditional pedagogical, reform pedagogical and transnumeration, findings were not as distinctive as with other tasks. Perhaps the large number of coded statements for the reform pedagogical category made it difficult to distinguish between categories. Another consideration was that ten students said they were unsure or needed more experience to describe how to tie material to the AP-Statistics exam. The Factory Task was also the first task on the survey, so perhaps respondents were unsure of how to respond to this question the first time. Further research may clarify some of the possible explanation (see *Reflections about the Limitations of this Study* section).

Factory Task belief responses.

Belief responses about the Factory Task (see Figure 4.17) show the comparative differences of how respondents felt about the Factory Task compared to the other seven tasks. An interesting trend with respect to the Factory Task was that this task ranked substantially higher than the mean response to the belief question "I have studied about this topic in a college class," with Alphas topping the three groups at rating the question 0.67 points higher than the mean response. This finding suggests that respondents could relate to the Factory task as containing topics that they had studied in classes, perhaps because of the histogram which was a familiar representation.

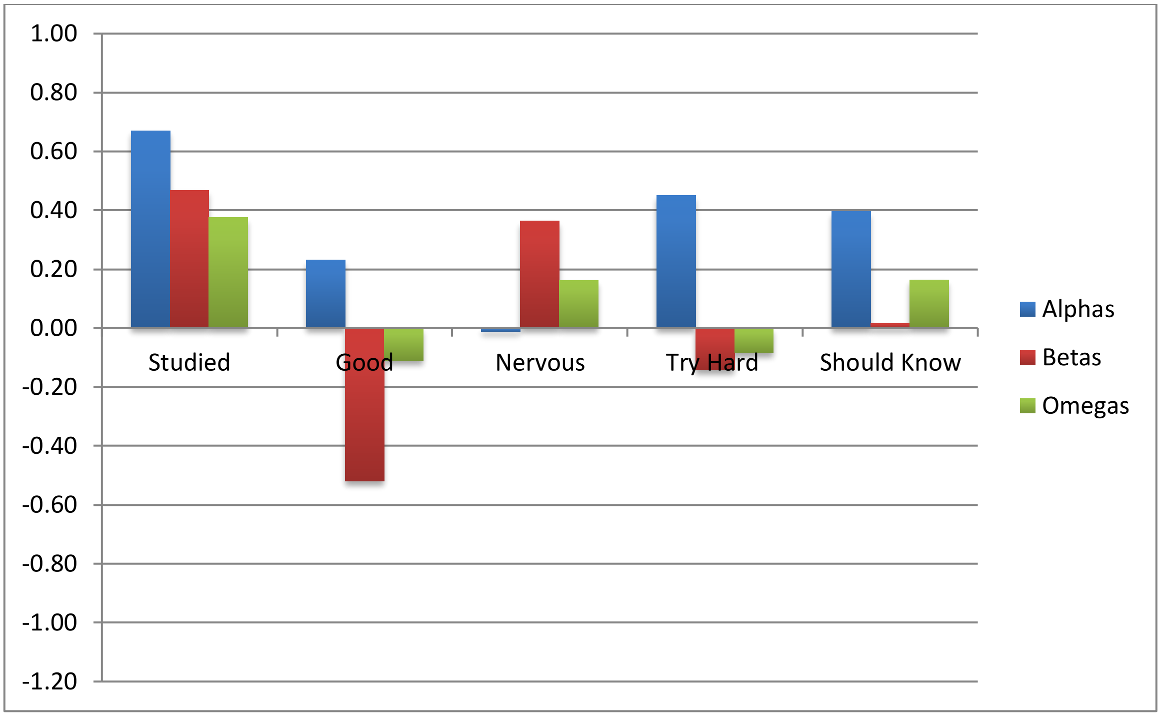


Figure 4.17. Comparisons of the mean response of belief questions to the Factory Task. Belief questions labeled above are as follows: Studied: This question is about a topic I have studied in a college class; Good: I am good at answering questions like this one; Nervous: I often feel nervous when I try to answer questions like this one; Try Hard: If I try hard, I can usually figure out questions like this one; Should Know: Secondary mathematics teachers should know how to answer this question.

Omega's belief question responses varied the least of the three groups across the five questions, with scores close to the mean response for each question. Alphas and Beta responses suggested different diverging feelings about the Factory Task. Recall that Betas had either passed AP-Statistics in high school or taken at least two college courses involving statistics or teaching mathematics. Omegas were all seniors or graduate mathematics students and had taken or enrolled in at least four courses about statistics or teaching mathematics. The first interesting finding about belief rankings was that even though Omegas had substantially more coursework then Betas, they ranked themselves as studying this topic just a bit worse (.38 average) than the Betas (.47 average).

When answering the second belief question "I am good at answering questions like this one," Betas believed they were not strong compared to other tasks scoring at -0.52, or just over a half a point lower on the Likert scale. Omegas on the other hand, remained very close to the mean scoring eleven-hundredths below. This finding aligned with responses from the third belief question about whether respondents were nervous in answering this question. Betas were more nervous than their average on this task (0.36 points above the mean for all seven tasks) while Omegas remained close to their mean at 0.16. Neither group felt strongly that they could answer this task if they tried hard (Betas = -0.14, Omegas = -0.08) or that secondary teachers should know this material (Betas = 0.02, Omegas = 0.16). Betas and Omegas for this study were two very different groups with regards to the qualitative interview findings. Interview findings suggested Omegas used transnumeration, statistical terminology and reasoning, along with pedagogical links to tasks at a far more consistent rate than Beta students. Perhaps a reason for the differences in belief findings for the Factory Task was that Betas believed they had studied the Factory Task content more, but were worse at answering this question and were more nervous about the content. Betas were in the thick of their statistical and pedagogical coursework in college, and were realizing they did not know as much about statistics as they thought (e.g., from low scores on exams). Whereas Omegas already had enough statistical experiences to realize that they had not studied the intricacies of statistics because of the limited amount of coursework time.

Locating places in teacher education programs where preservice teachers realize they *don't know everything* is critical to maximize a students' statistics learned, but also to reaffirming the importance of having a growth mindset (Dweck, 2008). With this we begin our segue between the robust results presented in Chapter 4, and a discussion of implications in Chapter 5.

Chapter 5: Findings and Implications

The primary focus of this dissertation was twofold: (i) to describe what preservice mathematics teachers know about graphical representations and transnumeration and (ii) to describe the extent to which preservice teachers were prepared to use graphical representations and transnumeration to help students understand statistical concepts. The previous chapter reported the data used to answer these two main research questions and the six more specific research questions. This chapter is divided into six sections: a summary of methods, a summary of results, answers to the six research questions, reflections about the limitations of the study, implications from and beyond this research study, and final remarks.

Summary of Methods

This research project was a multiple method, iterative process that included creating and revising tasks based on AP-Statistics assessment items, administering a survey that included the tasks to preservice mathematics teachers, and conducting paired task-based interviews to gain insight into survey results. Revising assessment items as exercises to research and support learning is a technique used by previous researchers (Groth, 2012; Watson, Callingham, & Nathan, 2009). For this dissertation study, creating and revising tasks began by repurposing released items from AP-Statistics exams into tasks that required or at least encouraged the use of transnumeration. Sixteen tasks were sent for review to three high school statistics teachers who had an average of 24 years of experience teaching mathematics and 16 years of experience specifically teaching statistics. One associate professor of statistics also reviewed the tasks. Reviewer feedback was used for task revisions and to decide on a set of tasks to use in a survey to study secondary preservice teacher knowledge. Specifically, reviewers were asked how closely the tasks targeted transnumeration, what teacher knowledge categories each task focused on, and whether interviews about the tasks would lead to better understanding of preservice teachers' ability to use transnumeration. Based on the input of the reviewers, eight of the original 16 tasks were selected for the final instrument. A series of belief questions about each task was also included to help identify preservice teachers beliefs about a task's value, motivation, and anxiety. Responses to the belief questions added to the story about why a respondent chose each answer. They also provided trends of beliefs about tasks compared to each other. For example, responses to the question, *should secondary teachers know this material* were compiled, averaged, and then were compared to other tasks to provide a perspective of what content secondary teachers thought was important. Similar to previous Statistical Knowledge of Teaching (SKT) research (Estrada & Batanero, 2008; Lancaster, 2007; Nasser, 2004) following-up a survey with interviews allowed respondents to justify answers and clarify beliefs responses.

The survey was administered to three classes of preservice teachers in a secondary mathematics program at a large university. Preservice teachers were classified into different three different levels of exposure to statistics and statistics teaching. As the GAISE report (Franklin, 2007) suggested, exposure to statistics tends to be more useful than age in determining statistical expertise. *Alphas*, the first group (n=13) were the least experienced in statistics and pedagogy. This group did not take AP-Stats in high school and took at-most one course in college that was either focused on statistics or teaching mathematics. A second group, nicknamed the *Betas* (n=11), had additional exposure to statistics and teaching mathematics. *Betas* had either passed AP-Statistics in high school (a year-long course), or they had finished at least two courses in college involving either statistics or teaching mathematics. The final group of preservice teachers, the *Omegas* (n=13), had the most experience in statistics and teaching mathematics. All *Omegas* were seniors or graduate students with a range of exposure to statistics and teaching mathematics. *Omegas* had taken or were enrolled in at least four courses about statistics or teaching mathematics. Because GPA is related to academic success (Chance, Wong, & Tintle, 2016), group GPAs were compared. There were no statistically significant differences in average GPA for the *Alphas*, *Betas* and *Omegas* for high school or college.

After the survey was administered across all three classes, the Salary Task was removed from the final analysis because no respondents answered, and could justify the correct response. For the remaining seven tasks, there were multiple correct answers to three tasks to which some respondents provided one but not all of the correct answers. Therefore answers were scored as *exactly correct*, *partially correct*, or *incorrect.* Distinguishing between exactly correct and partially correct answers helped to provide insight into the depth of preservice teachers' knowledge about graphical representations and transnumeration.

The number of exactly correct responses across Alpha, Beta, and Omega groups was relatively consistent (2.23, 2.18, 2.54) with each group correctly answering well-below half of the seven tasks. One possible reason for low scores was that tasks combined statistical content knowledge and pedagogical knowledge, which are unique and difficult knowledge types to cultivate for students (Haines, 2015; Watson, Callingham & Nathan, 2009). The average number of partially correct items increased across groups (Alphas 3.77, Betas 4.36, Omegas 5.00) although this difference was not statistically significant as determined by a one-way ANOVA ($F(2,34) = 1.509$, $p = 0.236$).

As previously noted, after each task, preservice teachers were asked five belief questions about the task. Responses options were on a seven point Likert scale (1-strongly disagree, 2-

disagree, 3-somewhat disagree, 4-neutral, 5-somewhat agree, 6-agree, 7-strongly agree). Responses were substantially different between groups for some of the questions. Alphas consistently stated they studied material less in college classes, responding with an average response of 2.64 compared to Beta (4.35) and Omega (4.70) students. These differences were strongly statistically significant as determined by one-way ANOVA ($F(2,34) = 97.813$, $p<.0001$). Some researchers view the use of a parametric procedure controversial because Likert scales use ordinal numbers, where values between numbers could not be viewed as equal (Norman, 2010). That being said, the significant ANOVA result indicated that, at very least there was a trend that respondents believed they were developing across the group classifications based on exposure. Alarmingly, there were some respondents that scored belief questions low *and* were relatively close to becoming an inservice teacher. For example, a majority of Alphas were actually juniors in college (7 out of 13) having only two years of courses left to develop their knowledge before becoming an inservice teacher. The combination of limited statistical exposure and a less than desired amount of time remaining in collegiate coursework creates a challenge to instruct preservice teachers best.

After survey completion, respondents were paired with other students they indicated they work well with to discuss responses in interviews. This purposeful sample encouraged communication between students and promoted discussion that helped answer research questions. Interview findings provided robust details about how respondents struggled to complete transnumeration tasks successfully. Interviews were flexibly structured to gather data for a wide-range of respondent's reflection styles (Watson, 2001), and coded using constant comparison techniques (Corbin & Strauss, 2014), which take into account findings from previous research. Three tasks were transcribed and reviewed to create clusters of comments (Miles &

Huberman, 1994), which provided intricate details about what respondents knew and how they used transnumeration in different ways. During interviews, preservice teachers discussed their responses to the survey tasks and were asked questions to evaluate knowledge of transnumeration.

Summary of Results

Evidence of using transnumerative thinking to visualize the reorganization of the dataset helped in responding correctly to the Tips Task (Appendix C). In brief, the Tips Task required respondents to look at a histogram of tip amounts and break histogram bars into data points to visualize how a mean and median would be influenced if a tip amount changed. As noted in the last chapter, 12 of the 16 respondents who showed evidence of using transnumerative thinking were completely correct (75%) whereas respondents who appeared to not use transnumerative thinking answered correctly only 11% of the time. This finding suggested visualizing "hidden" aspects in the dataset was a critical advancement to answering the Tips Task correctly. One of the Knowledge of Content and Students (KCS) interview questions asked respondents if a high school student would see anything different from the graphical representation than they see. Twenty-two of the 32 interviewees (69%) believed their translation of the graphical representation's information was the same as a high school student. Nineteen of these individuals (86%) justified their beliefs by talking about how the graph in the Tips Task was basic to interpret. Although respondents were learning how to teach secondary mathematics, the majority failed to highlight possible difficulties students could have in translating the graphical representation.

The next analyzed task was nicknamed the Fuel Task (Appendix D) and actually had two correct answer choices. The Fuel Task included three graphical representations, a boxplot, histogram and table containing the boxplot's five number summary. Respondents were asked how skewed the distribution was through using a ratio of $\frac{sample\ mean}{sample\ median}$. Only one Omega respondent and one Beta respondent chose both correct answers. Eight out of 13 Alphas (62%), eight out of 11 Betas (73%) and ten out of 13 (77%) Omegas were at least partially correct on the Fuel Task. Overall, respondents who stated that there was not a difference between secondary students knowledge and their own knowledge struggled on the Fuel Task, with only two out of six (33%) getting the task partially correct. Also, respondents who did not distinguish differences between high school students and themselves as teachers typically did not have a good grasp of the content. For example, Ava was a Beta who answered incorrectly on the Fuel Task. When asked if she saw something different than a high school student on this task Ava said, "I don't really think so, but I don't know. I think they would see like the same thing as what we are seeing." Ava's view of skewness within the Fuel Task was limited to her experience in AP-Statistics as she did not select either correct answer stating, "I don't really remember doing a ratio, so like sample mean over the median, I don't really remember that so I kind of eliminated 'A' and 'B' in my head." Respondents coded for pointing out differences between themselves and high school students did better than those who did not point out differences, with 15 of the 19 (79%) answering at least partially correct.

The final task that was transcribed and coded into constant comparison tables was the Factory Task. This task required respondents to navigate the limitations of a bimodal distribution of the diameter size of tortillas produced from multiple factory production lines. Only three out of 13 (23%) students in the Alpha group correctly completed the task. There were two common incorrect answers, both selected more often than the correct answer. One involved a stem-and-leaf plot (31%) and another focused on mean and median (38%). One possible explanation for respondent's choosing incorrectly so often was that, because they had little collegiate exposure to statistics and teaching mathematics, they relied on familiar topics (i.e., a mean). Only four of the 11 (36%) of Beta respondents were task at least partially correct, while six of the 13 (46%) Omega preservice teachers were at least partially correct. Beliefs about the Factory Task corresponded with the modest success with Alphas, Betas and Omegas hovering around neutral (average responses of 3.85, 3.91 and 4.38 respectively) on the statement "I am good at answering questions like this one."

Overall, the multi-method, iterative process employed in this research study provided a variety of opportunities for respondents to share their knowledge about transnumeration. There were clear differences between the three groups of preservice teachers. For example, four Omegas were at least partially correct on all tasks, but four different respondents were incorrect on more than half the tasks. In the interviews, some Omegas consistently discussed their Knowledge of Content and Students (KCS) and Knowledge of Content and Teaching (KCT), while others struggled to describe merely the Common Knowledge of the Content (CKC) used in answering a task. Because transnumeration is a type of statistical *thinking*, there are challenges to finding research methods that solicit respondents' thinking and gather meaningful data. A major advantage of this multi-method research design was it used different research methods and synthesized results together providing a more cohesive description of transnumerative thinking.

Answers to the Six Research Questions

The expanded research questions for this study focused on preservice teachers statistical knowledge, ability to use transnumeration, and pedagogical ability. This section provides eleven specific findings within the six research questions that were the focus of this study. Findings are numbered and listed in the final paragraphs of each research question section. Recall that types of knowledge, at the very least, have an influence on other types of knowledge, so findings discussed below often correlate between research questions. This perspective, along with a limited sample of students, resulted in a difficulty to describe explicit answers to research questions. Therefore, results to research questions below are not particularly *explicit*, but rather described as *trends* of the knowledge of participants that can be used to think about instruction and improvement within statistics education overall.

Research question 1a.

When answering research question 1a, "What statistical subject matter knowledge that preservice teachers display with graphical representations?" unique challenges to the age and development of preservice teachers complicate the results. Recall that the average number of tasks answered exactly correct in Alpha, Beta, and Omega groups (2.23, 2.18, 2.54) and partially correct responses (3.77, 4.36, 5.00), were not significant between groups by ANOVA. Perhaps the lack of distinction between groups was a result of institutional challenges in implementing statistics education coursework. The Statistics Education for Teachers (SET) report suggested there are differences in preparing teachers to teach statistics compared to other strands of mathematics coursework (Franklin, 2015). Preservice teachers often have different beliefs about life than inservice teachers while transitioning to adulthood. Supports for undergraduates are often limited based on university constraints of time and money, which results in limiting the number of credits a collegiate program requires. This limits the ease of adding courses for applicable, necessary courses with course descriptions focused on statistics, for other courses that have long been staples of a mathematics education degree. A side-affect of this limitation is since preservice teachers take an overload of coursework in traditional mathematics coursework, they typically use graphical representations to display mathematics rather than statistics. Results from the mathematics and statistics coursework misbalance are in some places profoundly obvious, while in other places more of a side effect of an overabundance in mathematics training. For example, other researchers found teachers feel comfortable with graphical displays, but struggled to use statistical language (Wessels, 2014).

Cole was a strong example of a respondent who felt comfortable with the material, but struggled with statistical language when he discussed his knowledge of skewness (see "Preservice teachers who completed AP-Statistics" in Chapter 4 for details). Cole claimed to understand that the dataset for the Fuel Task was skewed, but initially did not know how to describe what skewness was. As the conversation continued, Cole described skewness as "stuff not lining up as it should" on the histogram, and shifted the conversation back to pointing at the histogram to *show* me skewness. Cole's struggle to describe skewness in a graphical representation was a typical trend across Alphas, Betas, and Omegas. Some of Cole's struggle could be attributed to a limited exposure to varying graphical representation similar to how Cooper and Shore (2010) found that many inservice teachers struggled to highlight differences between types of graphs like bar graphs and histograms. This finding suggests that respondents' limited statistical language masked their ability to discuss knowledge of graphical representations and transnumeration with statistical terminology. Cole knew the histogram he was viewing was not typical, but could use words like *symmetrical*, *normal*, or even *distribution* or *histogram* to describe the skewed distribution.

One of my stronger mathematical participants during this research study was Sean, a graduate mathematics student who was taking pedagogical coursework to develop his capabilities as an aspiring future professor. Sean had limited exposure to boxplots throughout his mathematical studies. While interpreting a boxplot in the Fuel task he believed there was more data in the large quartile of the boxplot than the smaller one, a visually misleading but intuitive notion (Bakker, Biehler, Konold, 2004). Sean's graduate-level mathematical and reasoning capabilities often helped him work through tasks to correct answers, but for some of the less-intuitive statistical concepts Sean struggled especially to explain his answers.

Overall, there was comfort for most respondents in looking at the graphical representation portion of each task compared to scripted portions, which is a similar finding to Begg and Edwards (1999) who noted often respondents will use graphical representations to answer questions with unfamiliar text. What was unique about this dissertation's findings was respondents only felt comfortable with graphical representations if they were exposed to the type of graph previously. For example, the Real-Estate Task included a graphical representation with a cumulative probability distribution. Respondents often attempted to read the graph as a line graph, which they could relate to from their exposure in school. Respondents struggled to interpret the key aspects of the graph such as how a dependent variable was representing a percentage of the population. While interviewing about the Real-Estate Task, respondents used algebraic terminology like "x-axis," suggesting that they were familiar with functions on a coordinate plane but did not have experiences with statistics or probability distributions. In this manner respondents were operating with algebraic language where statistical language was needed to translate the context.

This research study's findings that summarize preservice teachers subject matter knowledge of graphical representations were:

1) Preservice teachers frequently did not have the statistical language to describe what was happening in each graphical representation.

2) Preservice teachers often tried to interpret the unfamiliar statistical aspects of graphical representations with their familiar mathematical content knowledge of graphical representations, which resulted in misinterpretations.

3) Preservice teachers were more comfortable and confident in interpreting the graphical representation portion of each task then the written description portion. Preservice teachers used the graphical representations similar to a crutch that they would lean on when struggling to answer a task.

Research question 1b.

Answering research question 1b, "What statistical subject matter knowledge do preservice teachers display with transnumeration between graphical representations?," was correlated with findings about research question 1a. Particularly, finding three discussed how preservice teachers used graphical representations similar to a crutch to interpret a task. This interpretation may have been, at least in part, done because respondents had limited exposure to statistics. Other researchers such as Watson (2001) found similar findings about the use of graphical representations when researching nine data and chance topics. Watson's results showed that teachers were *most confident* in teaching graphical representations compared to the other nine topics. However, because tasks in this research study were designed to incorporate transnumeration, respondents were required to look into graphical representations for information that was not visually obvious. This resulted in respondents having a false confidence in translating the information from graphical representations, or relying on intuitions in looking at a graphical representation and not drawing statistical conclusions.

The Tips Task was a great example of a task where respondents seemed to be over confident in their knowledge of the statistical subject matter, when in reality they struggled to interpret the statistical situation. The Tips Task was considered to be one of the easier tasks in the survey because its content covered mean and median, familiar topics to respondents. What was unique about the Tips Task was respondents were required to translate mean and median through a histogram, where many only had experiences of working with a list of numbers. Only 43% of respondents selected the correct answer yet respondents were more confident they were answering correctly than with other tasks. Specifically, Alphas, Betas and Omegas ranked belief questions about if they had studied this topic in class, were good at answering this question, and could answer this question if they tried hard above the mean across all tasks. Other researchers (González & Pinto, 2008) found secondary preservice teachers believed pedagogical knowledge, such as Curcio's stages of graph comprehension (1987), as unnecessary. Findings from the Tips task suggest perhaps a *false confidence* was influencing respondents to believe graphical representations were something a teacher can easily translate.

Other researchers, such as Rouan (2002) found teachers struggled to verbally extract information from graphical representations. This research study suggests some struggle to extract information from graphical representations is tied to a limited ability to use transnumerative thinking. Preservice teachers limited ability to use transnumeration along with their struggle to use statistical language resulted in misinterpretations of graphical representations. Wessels (2014) studied *inservice* teachers knowledge of variability and found they struggled with statistical language and transnumerating data with different kinds of representations after professional development opportunities. Perhaps this earlier exposure to transnumeration through tasks for preservice teachers will improve later professional development opportunity results. There is a clear need to reshape teachers beliefs about the importance of using different graphical representations and removing teacher's sense of false confidence that graphical representations are simple to interpret. Teachers need to be exposed to the power of highlighting different topics with different graphical representations.

Findings four and five describe overall trends in preservice teachers subject matter knowledge in using transnumeration between graphical representations:

4) A *false confidence* in interpreting the meaning of graphical representations often stunted the use of transnumeration. In other words, because respondents felt like they understood the graph, they felt no need to look for and use transnumeration between graphs.

5) Preservice teachers struggled to transnumerate between graphical representations often because transnumeration requires an unobvious depth in analyzing a graphical representation. A limited statistical vocabulary contributed to the inability to use transnumeration with datasets.

Research question 2a.

As explained in chapter 3, when expert reviewers reviewed the tasks answering research question 2a, "How do preservice teachers suggest they will use transnumeration as an inservice teacher?" reviewers suggested question 2a was going to be the most difficult question to investigate through the tasks. Research question 2a was only ranked as answerable by reviewers 8 total times across 64 reviews (four reviewers and 16 reviewed tasks), and no task was ranked as a task to help answer this research question by multiple reviewers. Perhaps the difficulty was because this research question was trying to predict the future, searching for answers on how preservice teachers viewed their eventual *inservice* use of transnumeration in the classroom. No matter what the reasons behind the difficulty in answering this question, the importance of thinking towards future instruction is a worthy and critical perspective to continue to link training material to future classrooms.

Research question 2a was also challenging to answer because many respondents struggled with statistical content knowledge, so explaining how to teach content as an inservice teacher added an additional level of difficulty. Other research studies have found teachers need to have CKC before other knowledge categories like SKC or PCK can really be developed (Noll, 2007; Watson, Callingham & Donne, 2008; Watson, Callingham & Nathan, 2009). Findings from this research study showed that respondents suggested to teach tasks with transnumeration, even when they were not using transnumeration themselves to solve tasks correctly. For example, the most common suggestion was to compare and contrast different graphical representation examples that highlight different statistical ideas. On the Tips Task, 12 respondents made statements that suggested as a teacher they would change the task to highlight properties or rearranging the dataset in another graph, but only six (50%) of these respondents answered the Tips Task correctly. This suggested preservice teachers had previous pedagogical experiences using transnumeration or at least rearranging graphical representations based on different scenarios, but could not use transnumeration themselves to solve the Tips Task. The lack of content knowledge in using transnumerative thinking resulted in a limited articulation about how respondents believed they could use transnumeration in their future classrooms, even though they believed using transnumeration was a high-quality pedagogical tactic.

Similar to Wessels (2014) findings, very few respondents articulated specific changes they would make to a graphical representation. Rather, they mentioned that rearranging a graphical representation is a pedagogical tactic that helps facilitate learning. Nick was one of the few examples of respondents who articulated their transnumerative thinking with statistical language (see the *Transnumeration with the median* section in chapter 4). For the Tips Task, Nick highlighted how the mean and median changed based on changes to the distribution. Nick suggested the pedagogical tactic of adding large tip amounts to the dataset multiple times so students could see the change in mean and lack of change in median. His description involved statistical terminology that suggested he use transnumeration, which was really impressive given many inservice teachers struggle with transnumeration.

Findings six and seven describe trends about how preservice teachers believed they would use transnumeration as an inservice teacher:

6) Preservice teachers suggested to teach content rearranging information which supports transnumerative thinking, even when preservice teachers generally struggled to use transnumeration to solve the tasks correctly.
7) Very few preservice teachers were able to identify changes they would make to a graphical representation to show how transnumeration highlighted specific statistical ideas.

Research question 2b.

A majority of expert reviewers rated both the Tips Task and the Factory Task as likely to provide information that could be used to answer research question 2b, "What will preservice teachers suggest students know about graphical representations? Transnumeration?" The majority of respondents for the Tips Task (69%) believed that their knowledge about the task was the same as a high school student. Perhaps some of this belief might be credited to the Tips

Task using the familiar pre-college topics of mean and median. With only eight preservice teachers making statements that were coded for highlighting differences between high school students and themselves, and only two of these statements specifically being differences about the graphical representation, preservice teachers rarely articulated important attributes in graphical representations. The limited ability of teachers to describe information from a graphical representation was consistent with Rouan's (2002) findings about teachers limited statistical CKC. What was new from this research study was a clearer picture of *why* preservice teachers struggle to interpret some graphical representations. On the Tips Task 73% of respondents correctly believed that both mean and median would increase (by answering C, D, or E), 46% answered at least partially correct. However, only 25% of interviewees were able to specifically note a difference in what knew about the graphical representation compared to what they believed a high school student knew, and only 6% highlighted a specific difference about the content in the graphical representations. There was a clear diminishing trend in preservice teachers' ability to justify a correct answer with the level of detail that would be expected when teaching students.

For the Factory Task, 15 respondents (41%) statements were coded into a *no difference* category between their knowledge and a high school student's knowledge (see Table 4.41). Respondents statements were then coded into two subcategories: (i) a general no difference subcategory and (ii) a subcategory focused on the idea that histograms are not a difficult topic to learn. Although the Factory Task was rated as a more difficult task for respondents then the Tips Task, it was interesting that such a large percentage of respondents did not think there was a knowledge difference in translating the graphical representation compared to a high school student. Of the 14 students that believed there was a difference between their knowledge of graphical representations and a high school student's knowledge of graphical representations, four subcategories were created including a subcategory in which six respondents mentioned histogram information or transnumeration was the difference. Their was a clear pattern of personal beliefs influencing how to interpret graphical representations or the use of transnumeration to solve tasks, similar González's (2016) finding that knowledge types and beliefs are both important to develop in teachers. For the Factory Task, when respondents believed they had different CKC than high school students, they offered more detailed reasons than those respondents who believed their knowledge was the same as high school students.

There are several possible reasons for why respondents often did not believe their knowledge was different from a high school student. Some respondents did not have the content knowledge necessary to complete some tasks, which linked to their lack of confidence in describing what was happening in a graphical representation. In other words, graphical comprehension (Friel, Curcio, & Bright, 2001) was limited. Without necessary statistical language or graphical comprehension to describe tasks, respondents are correct that they lack deep understanding of graphical representations, a necessity for teaching material well. Another reason why respondents did not believe their knowledge was different from a high school student's knowledge was similar to the findings of González and Pinto (2008); respondents viewed reading graphical representations as a simplistic task that would not confuse students. Together, a limited graphical comprehension and statistical vocabulary with beliefs that their future students would already have necessary graphical abilities to translate graphical representations creates a difficult situation for college instructors in the root of the deficiency is very different. Preservice teachers without the necessary CKC need supports to develop their *knowledge*, but preservice teachers that believe graphical representations are easy to translate need to correct their *beliefs*.

Findings eight and nine describe trends about what preservice teachers suggested students know about transnumeration:

8) There was a clear diminishing trend in preservice teachers' ability to justify correct answers with the level of detail that would be expected when teaching students. This suggests preservice teachers would struggle to answer student questions, or clarify details in student ideas.
9) Preservice teachers without the necessary CKC need supports to develop their *knowledge*, but preservice teachers that believe graphical representations are easy to translate need to correct their *beliefs*.

Research question 2c.

When answering research question 2c, "What pedagogical tactics will preservice teachers suggest to teach concepts in graphical representations?" respondents typically did not suggest teaching methods similar to options in the task. Instead respondents used their instincts of how to teach statistics within tasks that was often related to other factors outside of the task. Perhaps this meant that even though tasks can be written in a manner to influence pedagogical choices, instinctual pedagogical choices can over-ride the given options. Some trends in the pedagogical choices offered were bridging-off previous responses to KCS or KCT questions or aligning responses to their other paired-interviewee, however responses often seemed to be driven by beliefs because they were developed in a variety of directions.

A good example of respondents not utilizing answer options was the Tips Task (Appendix C) in which 17 respondents answers were at least partially correct choosing answer choice “E” which involved having the student cut out histogram bars, tape them together by length and fold the entire length to find the median of the dataset. Answer “E” is a clear pedagogical tactic, explicit to the content in the Tips Task. The tactic came from a practicing AP-Stats teacher who uses it to teach about medians in histograms. There are great reasons to merely use the given pedagogical tactic for this situation, but most respondents did *not* use it! Only five respondents answered responded to questions about how to teach Tips Task material with “answer E,” in which two of the five respondents actually did not answer “E” during the survey or interview! This leaves three respondents out of 17 (18%) who answered “E” on the survey that believed the best pedagogical choice was presented as an answer choice. One explanation may be similar to findings that teachers needed to have other knowledge categories like SCK before PCK topics can be cultivated (Watson, et al., 2008; 2009). Respondents without SCK about how material in the Tips Task could be used to teach were left without the connections to link answer “E” as a pertinent pedagogical tactic.

Pedagogical suggestions spanned a wide-range of possibilities because every respondent had their own personal beliefs already about what teaching should be. The number of main categorical trends for the Tips Task (see Table 4.14); the Fuel Task (see Table 4.28) and the Factory Task (see Table 4.42) varied. Perhaps this variation suggests that even though respondents were in the same collegiate program they did not really view pedagogical choices as something that could be proceduralized or cultivated in a best-practices direction. Often categories used vernacular that is generally accepted as good pedagogical choices like the *real-life applications* subcategory, however respondents provided descriptions of what a real-life application in the was often underdeveloped. A clear result of interviews was each respondent's beliefs would play an influential role in how they would teach graphical representations in their future classroom.

Finding ten describes trends about the pedagogical tactics preservice teachers suggested to teach concepts in graphical representations:

10) Preservice teachers displayed limited Specialized Knowledge of the Content by suggesting a wide-range of pedagogical tactics that were different than answer choices. Suggestions were at least partially driven by their beliefs about the best way to answer a question.

Research question 2d.

When respondents were asked "What teaching tactics they would use to teach the material to a class of AP-Stats students to help prepare students for the exam?" respondents often struggled to offer task specific suggestions. Respondents did articulate general pedagogical tactics to prepare students for standardized tests that their AP-Statistics teachers modeled or other AP-teachers utilized, similar to one of the three APS-PCK knowledge categories Haines (2015) suggested were needed to teach AP-Statistics. Examples of general pedagogical tactics were having students complete released exam problems, and grading old exam problems with similar standards to how AP-graders grade tests.

Respondents frequently described a similar situation when responding to questions six and seven on the interview protocol (see Figure 3.16). Question seven was about how respondents would prepare students for the AP-exam and questions six was about how respondents would teach this topic to a classroom full of twenty students. The intention for distinguishing between these two questions was to evaluate what specific statistical pedagogical tactics respondents believed would benefit students beyond high-stakes assessment expectations. However, most failed to differentiate between pedagogical practices and AP-Statistics exam preparation. For example, on the Tips Task, 18 respondents, pedagogical suggestions were coded in the same category. One benefit of the consecutive nature of these two questions was that respondents had a chance to rethink questions, and often went into more depth on their answer to the sixth protocol question as they provided their answer to the seventh protocol question. This complicated the analysis of this research question, which was noted in constant comparison tables like the table found for the Fuel Task (see Table 4.29).

One critical finding about respondents KCT was that if respondents were able to utilize specific statistical vocabulary during interviews tied to graphical representations, they were more likely to answer the task correct. This finding is similar to research indicating that there is a progression of knowledge types for many preservice teachers, beginning with CKC (Noll, 2007; Watson, Callingham & Donne, 2008; Watson, Callingham & Nathan, 2009). One challenge of looking at only the survey results and multiple-choice responses was the limited amount of data gathered about the knowledge of participants. The multi-method nature of this study with written (survey) responses and oral (interview) communication distinguished respondents' different levels of understanding graphical representations and transnumeration. This suggests that to analyze preservice teachers KCT of graphical representations and transnumeration, a variety of data sources were necessary.

Finding eleven describes trends of teaching tactics preservice teachers would use to teach the material to a class of AP-Statistics students to help prepare students for the exam:

11) Preservice teachers typically offered general suggestions about teaching AP-Statistics, not task or content specific suggestions.

Reflections About the Limitations of the Study

When proposing this research study, I knew I would need multiple-methods to answer my research questions about transnumeration. Most quantitative research studies become engrossed in the numbers and can lose the true meaning they are trying to measure, while qualitative research can lose the descriptive statistics that provide an overall perspective on research variables. This multi-method research study attempted to utilize both research styles to tell the story of how one university's preservice teachers learned and thought about transnumeration during their secondary mathematics program. The purpose of this section is to highlight some of the strengths and weaknesses of this research project and justify some of the decisions that were made along the research path.

One specific challenge in this study was a limited sample. With only 37 students from one university's secondary mathematics teacher preparation program readily available, the results are based solely on the unique set of teachers that were available at the time of the study. One method that strengthened the study, and thus helped to make the results as robust as possible given the sample, was utilizing three high school statistics teachers and a university statistician to rate and critique tasks before implementing them with respondents. The critiques of tasks by the teachers and statistician helped assure that the survey and interviews would focus on what experts view as the phenomenon of transnumeration. Another technique this research study employed was writing substantially more tasks than needed and the selecting the most appropriate subset of those tasks for the study. Of the 30 tasks that were at least partially developed and the 16 sent for expert review, only seven were included in the final analyses.

The nature of transnumeration itself was another difficult challenge of this research study. Because transnumeration is a type of thinking (Wild & Pfannkuch, 1999), tasks had to be written in a manner to foster transnumerative thinking and promote dialogue that allowed respondents to describe their thinking. Previous research has already documented the challenges in researching both content and pedagogical knowledge (Hill, Ball, & Schilling, 2008). Additionally, we know from research that factors such as the length of question and answer choices or the position of correct answer choices (Rodway, Schepman, & Thoma, 2016) can influence a response choice and thus tasks had to be designed so respondents could understand tasks with a short word-count. The expert review phase helped limit the ambiguity of having an abstract-topic like transnumeration woven together with many different contextual elements required in the field of teaching.

As the lone researcher, the generated constant comparison tables were bound by the translation of respondent's words into my language, and did not have the opportunity to stand across a team of researchers for interpretation of results. The data and conclusions from the comparison tables were, however, discussed with the advisor for this research. Using clusters of comments (Miles & Huberman, 1994) along with searching for meaning condensations (Kvale, 1996) helped in summarizing the data. Constant comparison tables were organized by interview questions, which were specifically designed to research Common Knowledge of the Content (CKC), Specialized Knowledge of the Content (SKC), Knowledge of Content and Students (KCS), and Knowledge of Content and Teaching (KCT) categories. Constant comparison tables were then recoded across-interview questions. For example, the seventh intereview question was designed to facilitate a discussion about respondents KCT, but respondents may have provided further information about their CKC while answering this question. This CKC information was added to CKC constant comparison tables for a better depiction of respondent knowledge. Coding statements across-interview questions ensured CKC, SKC, KCS, and KCT categories included all knowledge respondents provided regardless of when during the interview relevant comments were made. The open-ended nature of the interviews created challenges in that each respondent was answering a somewhat different set of questions. However, without having flexibility during interviews, the ability for respondents to share as much information as possible about their knowledge of transnumeration would not be possible..

One advantage of being the lone researcher was that I conducted all the interviews myself. I understood the context of statements and could remember statements that respondents made during interviews that would require further investigation. I had field notes to reflect on different questions and began to compile trends on how respondents were answering the different tasks, even before full transcriptions of interviews were available. Occasionally, respondents described their use of transnumeration on a task with a very limited or informal vocabulary. Being the lone researcher I knew what was discussed in interviews very well and could recall the context of statements to help describe their meaning (Kvale, 1996).

Implications From and Beyond this Research Study

The following reflections are provided to help professionals in key influential positions consider how this research study can be used to help improve statistics education. Sections are broken-down into a section addressed to parents, to gatekeepers or those in charge of structuring children's early statistical experiences (e.g., mentors, and math coaches) along with a section addressed specifically to statistics educators. Sections are written to communicate implications of this research to different audiences, tailored by their different use the language and word choice. Therefore, a section written *To parents* offers general suggestions that may reach beyond this research about changing the general educational influences in statistics for children. In contrast the *To gatekeepers* section provides more explicit links to research and ties more closely to the application of this research that gatekeepers themselves can translate in their work to impact others (i.e., using tasks to facilitate professional development experiences). Finally, the section written *to statistics educators* provided detailed information on the impact of this research study and how other research might utilize results to further the field of statistics education. Although many of the ideas throughout the following sections are not all new discoveries from this research study, they do provide pragmatic applications for important stakeholders.

To parents.

I had a conversation with an undergraduate student excited about a new class being offered at my university focused on Data Science. The undergraduate was excited about the data-mining coursework he was taking, saying "they say, this is the future of business." I tend to believe strongly that using data to make business decisions has been for the future of business for centuries, albeit in very different forms. The way we gather and use data is rapidly changing, but when it comes down to it, smart businesses are still finding and using data to make decisions. As you think about the influence you can have on your child and their ability to learn statistics, I feel as though very few parents will argue that there is not a purpose to learning statistics. Applications of statistics are embedded throughout most professional fields today. Parents are responsible for early statistical experiences for their children, which should be open-ended, curiosity driven, flexible and fun statistical experiences. If possible tailor statistical experiences to your child's interests. A few simple thoughts to improve your child's future and support statistical futures are:

(i) *Ask questions and encourage great questions* – A current challenge to parents is to slow-down and take the time to not only ask great questions, but to cultivate great questions from their children. Opportunities with statistics are endless, finding datasets about different topics are readily availably now online. Graphical representations often incorporate intuitive concepts that scale down well into the early childhood grade levels. Typically, the question is not whether a young child can have great, statistical experiences at a young age, but rather will we take enough time to ask questions and use data sets that are age appropriate.

(ii) *Do not close door, build homes* – I recently had a conversation with a parent of a 4th grade student who talked about how his son believed traveling backwards in time was impossible because our universe is always expanding and earth would not be at the same location as it is right now. A topic that typically would be viewed as out of the mental comprehension of a 4th grader, the parent was very proud of his child's mental stride, and should be proud that he did not close the door on this topic of time-travel even with advanced nature of the topic. The challenge in these kinds of learning situations for parents is to go a step further for your child and find resources that can *further* challenge students with topics and materials that will continue to stimulate thoughts about a topic. I encourage parents to think about the metaphor of closing doors in a house compared to building more rooms on. Be careful to focus your children in so much (i.e.

closing doors) that you stunt creativity. Instead, pursue understanding the interests of your child, or keep the doors open and walk into those rooms. Help your child develop their interests like you would help them organize their bedroom. If possible, build additional rooms as your child's interests develop and help you child see connections between their interests. You will end up with a child that thinks flexibly, makes academic connections, and can participate in unique conversations across disciplines.

(iii) *Find statistics in puzzles around YOU* – I have fond memories of solving 100+ piece puzzles with my father from a very young age. I think it was one way my father, not knowingly, cultivated my own interest in spatial reasoning and in some ways starting this research topic of studying transnumeration. Although puzzles clearly had an impact on my own mathematical ability, perhaps as important was the *time* and *encouragement* I received from my father as we solved the puzzle. Opportunities to work on mathematical and statistical puzzles are flourishing, particularly with technological resources. However, often we are tempted as parents to get children started with technology and no longer participate with children as they explore topics in technology. Sort through resources to try and incorporate at least a few that cultivate your child's ability.

To gatekeepers.

I was recently rereading Jo Boaler's famous book, *What's Math Got to Do With It?* In the last chapter, Boaler wraps up her narrative by discussing twelve pragmatic actions that can be taken to improve students' mathematical experiences. She mentions actions like never telling a child they are smart but rather focusing on what the child did that was smart, and encouraging parents to avoid telling children they will be bad at mathematics like their parents were during school years. She provides practical suggestions about what we could all do to improve the development of children through mathematics. Four of the twelve actions Boaler described began with the word “encourage.” I cannot help but think that the statistics education community, more than anything else right now, needs to be a group of encouragers to our future teachers. During interviews, frequently secondary teachers described their answers to tasks as educated guesses based on limited knowledge. The majority of Alpha’s were juniors in college, and responded to the belief question, “This question is about a topic I have studied in a college class” an average of 2.64/7 compared to Betas (4.35) and Omegas (4.70; see Table 4.4). I received so many justifications about a limited exposure to statistics coursework, or pedagogical experiences that were far from students *doing statistics*. Over and over, I found myself encouraging preservice teachers to explore statistics further. I made a constant effort to try and help respondents believe that their abilities to teach were only beginning and that they had so much potential if they did not look at tasks for *right* and *wrong* answers, but rather opportunities to grow while completing tasks from mistakes. Some respondents clearly had growth mindsets during interviews (Dweck, 2006), searching for *why* responses were correct or incorrect.

As these aspiring secondary mathematics teachers continue their journey, they need to be supported in learning how to use a growth mindset. This dissertation research clearly showed that many secondary mathematics teachers will graduate college and begin teachers while still needing to learn more statistics. This study also showed that many preservice teachers were cultivating a fixed mindset, trying to merely survive collegiate coursework instead of learn material and how to apply material to their future classrooms. Therefore, perhaps one of the most important tasks for gatekeepers is to cultivate a shift from fixed to growth mindsets.

Using tasks with teachers.

One way gatekeepers can begin the conversation with teachers about teacher knowledge is to use tasks to discuss pedagogical classroom choices. Using a task as a professional development support or teacher training prompt can help respondents rethink how to justify classroom activities, their consequences to student learning, and still revisit many content aspects of statistics that need to be refreshed. These transnumeration tasks are tools for mathematics teacher educators to highlight statistics and help support the development of preservice teachers' SKT. Task experiences can strengthen beliefs in how to teach through classroom scenarios preservice teachers could encounter in the field. Because research in teaching statistics calls for an extremely student-based pedagogical style that is new to many teachers (Franklin et al., 2015), exposure to some pedagogical techniques through prompts in the collegiate classroom can be the difference in the change for a preservice teachers' K-12 classroom.

Results from this dissertation showed that having respondents work in pairs and discuss tasks, more often than not, resulted in respondents changing answers to correct answers than incorrect answers (see Table 4.5). Jeremy and Eve (see section "Interview discussions promote change") had a discussion during their interview that elicited new knowledge and resulted in them changing answers to the correct answer. Given the primary purpose of interviews was to research knowledge and not cultivate knowledge, I believe facilitated discussions could be even more productive in statistics or methods classes.

Tasks naturally provide teachers with the opportunity to explain the *why* behind the KCS in the task. Results from this study showed when respondents believed their knowledge was different than a high-school student *and* could articulate that difference, they were more likely to respond correctly to both the Fuel Task and Factory Tasks. The percentage correct who stated there was not a difference between their knowledge and student knowledge was 33% for the Fuel Task and 27% for the Factory task. Compared to those respondents who believed there was a difference between their knowledge and student knowledge for the Fuel Task (79%) and the Factory Task (50%), there is quite a difference between these groups. Perhaps this is because preservice teachers who believe there was a difference in their knowledge and student knowledge are thinking forward on a consistent basis already about their future as inservice teachers.

The Tips Task did not follow the same pattern having 45% of respondents answering correct when stating there was not a difference between their knowledge and student knowledge compared to 44% answering correct stating there was a difference between their knowledge and student knowledge. Perhaps the Tips Task results did distinguish between groups similarly because a larger portion of respondents (22/31) believed there was no difference in their knowledge and students knowledge compared to the Fuel Task (6/31) and Factory Task (15/31). The Tips Task was a more familiar content topic, and belief responses were more positive than on the Fuel Task and Factory Task, which may have played a role in these results.

There is plenty of research that shows that teacher beliefs are a critical influence on pedagogical choices, and that in order to change classrooms consistently beliefs need to change as well (Philipp, 2007; Pierce & Chick, 2011). I encourage gatekeepers to make time for the conversations with new teachers about how it is possible to teach statistics in a manner that cultivates creativity in children.

To statistics educators.

The culmination of my statistics experiences to from taking numerous collegiate statistics courses, teaching K-12 students and preservice teachers, and conducting an early-doctoral research project focused on probability is this dissertation study about *transnumeration*. Every statistics educator makes a contribution; some contributions are built off other statistics educators trailblazing for the field, while many statistics educators make their impact educating high school and college students about statistics. I believe this research study can open doorways for formal research studies at collegiate institutions across the country, but also for small-scale action research projects in high school classrooms across the country. The topic of transnumeration is specific and unique idea of thinking, but it spans across many different possible graphical representations and topics. Researching beliefs about transnumeration can take many forms as well, with a focus on statistical content, or how one translates a graphical representation. There are many opportunities to researching and understanding more about how we use the concept of transnumeration.

Work involving transnumeration is emerging across the research field of statistics education because of the advantages technology provides in quickly reshaping distribution. Recent research findings suggest teachers have different abilities and comfort levels in using transnumeration. Lee et al., (2014) found that 72% of the teachers believed using one graphical representation was enough to display information, a finding that supports this research's findings of preservice teachers struggling to flexibly work between graphical representations or notice hidden, unobvious attributes of graphical representations. Many teachers *believe* there is enough information in one representation for students and will limit classroom opportunities for students to use transnumeration between different types of graphical representations.

Graphical representations traditionally are thought to be critical to displaying information in many different fields such as science, statistics and mathematical fields. However non-traditional subjects like language arts are seeing an emergence of using graphical representations to develop reading comprehension. Graphical representations place in curriculum is now *a little bit everywhere,* which results in teachers using graphical representations frequently, but teaching intricacies of graphical represents *rarely.* For example, graphical representation standards are found across many different domains in the common core standards for mathematics, but graphical representations are not identified as a specific topic of study by themselves. The *Measurement and Data* strand begins second grade standards have students drawing picture and bar-graphs in different categories (CCSSM2.MD.D.9) and has fourth and fifth grade standards teaching with using line plots (CCSSM5.MD.B.2), but ends from fifth grade till high school. *Operations & Algebraic Thinking* standards in third grade call for the use of arrays, drawings and equations when teaching multiplication and division (CCSSM.3.OA.A.3), but still do not explicitly call for comparisons between different graphical representations or for analyzing a choice between displaying material with different graphs. In fact, the most impactful standards for analyzing graphs may be with functions in eighth grade when the standards call for modeling relationships between quantities and "describing *qualitatively* the functional relationship between two quantities by analyzing a graph." (CCSSM.8.F.B.4). Are we surprized that respondents in this study *heavily relied* on experiences from Algebra? Even when looking at sixth grade proportional reasoning standards, which are strongly linked to some statistical topics, standards require Algebraic ideas about using a coordinate plane to understand negative integers. Perhaps it is time to incorporate graphical representations as a domain topic within the common core standards. After all visuals are a universal language for our, more than ever, globally connected world. I am not suggesting graphical representations are not presented in CCSSM, but rather graphical representations importance to mathematics and statistics is much more important than the allotted attention. For example, teachers need to be using pedagogical tactics that demand reasoning through selecting the most important graphical representation for a situation., Students need to be able to draw their own visual display of data and compare similarities and differences with other students and in many cases expose students to new types of statistical thinking like transnumeration.

We know interpreting graphical representations is important, but there is clear evidence that displaying information through different graphical representations is changing with technological advances. It is possible the most important graphs for children to understand as professionals are not closely linked to a coordinate plain. Many students are being trained with a lack of flexibility in identifying key factors in graphical representations (i.e. labels, identifiers) and more importantly a stagnant-style of thinking about better and worse representations across grade levels. We should not be surprised at findings that the general public struggles to interpret graphical representations involving statistics in the news, nor the different calls to improve the public's statistical literacy (Engle, 2017; Watson 2013). There is a gap in how we are teaching graphical representations in our standards! I believe the answer may be to explicitly include graphical representations as one of the domains into the common core standards as a ninth mathematical practice. On the surface, this request seems to be a rather enormous suggestion to which there are valid arguments against it like, the bulkier standards get the harder they will be to teach. But there is a clear importance of graphical representations across many different classes and professional fields for students. It seems only intuitive to train students to explicitly on interpreting different types of graphical representations with versatility and flexibility to equip students for 21st century careers.

Future research directions.

I believe the opportunity to research knowledge and use of transnumeration calls for more research across grade levels. Recall Sean was a graduate mathematics student working towards eventually teaching in college. Sean did not know that smaller boxplot areas implied the data was more condensed (see section *A misunderstanding about boxplots*). This was a great example that there are statistical concepts that even advanced graduate students need to be trained on. Other researchers have documented similar challenges for learners when interpreting boxplots (Bakker, Biehler, Konold, 2004). The topic of building a boxplot alone is often introduced in middle school grades, but as seen with Sean, still has a place with context to providing challenging questions to graduate mathematics students. Sean's unique position as a graduate student peaked my interest in wondering how graduate mathematics students use transnumeration and more generally how graduate students analyze graphical representations.

Transnumeration could also be scaled down and researched in younger grades. For example, elementary students create bar graphs and line plots, but I wonder how flexible their interpretation of these beginning data displays is? I observed Kindergarten lessons where students make bar graphs of the different colored Skittles they were given, I wonder after this early experience with bar graphs what noticeable differences in describing data students make? The question is not whether it is possible to study transnumeration across grade levels, but rather how other statistics educators will think to study the topic of transnumeration.

A different direction the statistics education research community could pursue is how to foster a growth mindset (Dweck, 2006) for secondary mathematics preservice teachers while using discussion tasks like those developed in this research study. Although it was beyond the scope of this research study, I noticed that some respondents enjoyed the interviews conversations searching for *why* each correct and incorrect answer was included in questions. Other respondents attempted to avoid conversations, or limited their word-choice that would display a lack of knowledge in answering tasks. Research needs to be done on the influence of growth and fixed mindsets across different topics: learning statistics, interpreting graphical representations and using transnumeration. I believe researching how to change mindsets of preservice who are operating with a fixed mindset could be critical in cultivating long-term, highly-qualified secondary teachers.

Finally, the data from this study shows that many preservice teachers are graduating across the country with limited preparation to teach statistics with graphical representations and there are limitations in expanding statistics coursework. Recall, ANOVA results showed no significance between groups (Alpha, Beta, and Omega) for the average number of tasks answered exactly correct (2.23, 2.18, and 2.54 respectively) or the partially correct responses (3.77, 4.36, and 5.00 respectively). Further research using a larger sample from universities across the country could provide a much better description about if limited development in preservice teacher knowledge of transnumeration is a national trend. Whether it was a limited exposure to content or pedagogical tactics, many respondents in this study showed gaps in knowledge that will be filled in during teaching experience at the sacrifice of their 1st or 2nd year students. GAISE provided a clear pathway for students to gain statistical experiences through its four-step statistical investigation process of: (i)Formulate Questions, (ii) Collect Data, (iii)

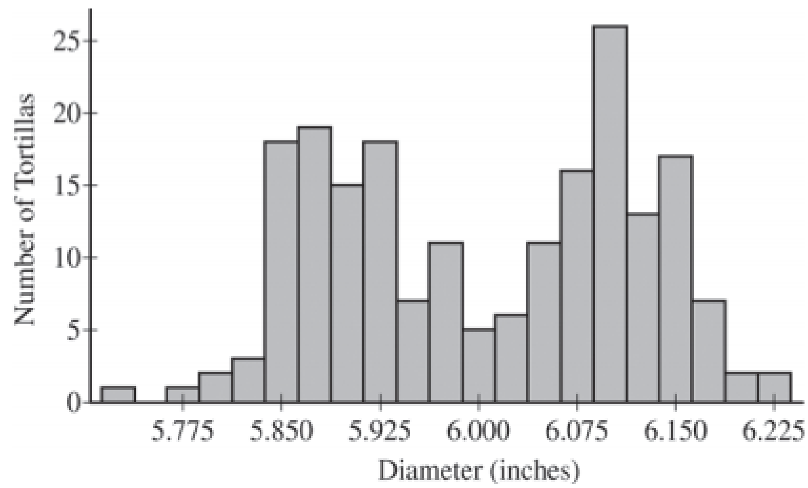
Analyze Data and (iv) Interpret Results. Analyzing data requires students to "select appropriate graphical and numerical methods" (Franklin, 2007, p. 11), which places a critical importance on understanding graphical representations. Unfortunately findings from this research study showed the limitations of preservice teachers knowledge of different types of graphical representations, especially of unique representations. A larger sample would provide information on whether this limitation was unique to one university, or a national limitation exists perhaps in our curriculum layout. How readily should we expect a new teacher to provide *flexible* experiences for students to use different kinds of representations if they were limited to *repetitive* experiences of viewing statistical distributions through only histograms and boxplots?

Final Remarks

Just over twenty year ago, Susan Friel called for improvement in teacher knowledge of graphs and how to teach graphs (Friel,1997), noting than many practicing teachers did not have the opportunity to work with statistics and data analysis in school. Over the past few decades, opportunities for teachers to work with statistics have changed dramatically. Statistics education is now prominent in school districts with many offering statistics a separate course for high school students (Pierson, 2016). Many districts have groups of teachers trying to infuse the curriculum with statistical experiences across middle and high school to support students' statistical development across different levels (Franklin, 2007). The importance of secondary mathematics teachers understanding statistics was once optional, but now we are at a critical junction where every teacher should be able to facilitate a discussion involving data. Teachers need the statistical literacy, reasoning and thinking (Ben-Zvi & Garfield, 2005) to construct meaningful classroom experiences for students. Given the public views teaching statistics as important, Friel's call to improvement teacher knowledge about graphs is more achievable than ever before. As researchers, the challenge now becomes what can we do to improve the knowledge and use topics like transnumeration in graphs in K-12 education?

Recently, the American Statistical Association began to release a monthly visual through the New York Times called *What's Going on in This Graph* in an attempt to fuel the statistical conversations about graphs (NY Times, 2017). Hundreds of educators are commenting on the visuals created and how they can be used with students to represent explorations of datasets, searching for ways to describe phenomenon hidden from the common eye. The professional development of secondary teachers is happening more than ever, bottled into a user-friendly conversation across state lines and even national borders. Conversations are happening. As a research community we merely need to direct the conversation.

Appendix A: The Factory Task's Expert Reviewer Feedback Template

Factory Task	
Task	Peggy is a high school student in your class working on the following problem: A factory's goal is to produce tortillas at 6 inches in diameter. The quality control workers in the factory take a random sample of 200 tortillas across multiple production lines. Results show a sample mean of 6.02 with the following distribution:  Peggy describes the factory's performance saying "The factory is doing a great job of meeting their goal. The sample mean is only .02 inches away from the population mean goal." What should you do next as Peggy's teacher to further her understanding of the factor's performance?
Possible Answer Choices	A. The factory's performance is outstanding being only .02 inches away from the mean. As her teacher, I would agree with Peggy and ask if she thinks the company should take another random sample of 200 tortillas to verify results. B. As Peggy's teacher, I would try to highlight why a histogram is used to represent data. Have Peggy create a Stem and Leaf plot by hand to see how the histogram above was developed. C. Calculate the mean and median of the data set, and then ask Peggy to compare how the two measures relate to each other. Finish by highlighting the critical differences between the mean and median. D. Have Peggy draw an normally distributed histogram and says it's mean is also only .02 inches away from 6 inches. Then ask which production line in the factory has better performance based on the two distributions.

	E. Peggy should change the scale on the y-axis from an interval of 5 to 1 to get a more precise histogram of the amount of tortillas with each diameter range. Then ask Peggy to explain if the sample mean changed at all given the more detailed histogram.
Answer Key	Answer: D
Rationale for Key Intended Purpose of correct answer AND distractors	A. This distractor answer focused on Peggy's excitement for focusing on a popular measure in the mean. Often the mean is consider a great description of any data set, in the case of a bimodal graph it does a poor job of describing many of the data sets attributes. Peggy should be concerned that 2 different places were slightly off in producing tortillas. B. Using a stem and leaf plot could lead Peggy to seeing that the vast majority of numbers vary far away from 6, even though the sample mean is very close to 6, therefore this answer's content knowledge is reasonable. In other words, a stem and leaf plot can be used to piece together a histogram and get a feel for its center, spread, and context of the problem. However having a student create a stem-and-leaf plot by hand with 200 numbers will be too lengthy of a task logistically for a teacher to implement with students, making this answer the wrong choice with pedagogical considerations in mind. C. This distractor focuses on measures of central tendency to add more attention to that concept along with answer A. D. This is the correct answer. Asking for a student to visualize a normal distribution should focus the student on the differences in range for the samples, and provide an opportunity to talk about why a bimodal distribution is occurring, or perhaps that there are 2 different places that have errors that could be adjusted to get a 6inch diameter. E. This distractor is true in that more details would be possible when changing the scale of the y-axis from 5 to 1, but this move would not make a difference in the graph forming a bimodal distribution.

1. Recall transnumeration is a type of statistical thinking. How much does someone completing this task have to use transnumeration?

__ Transnumeration is required to answer this task.
__ Transnumeration is helpful, but not required to answer this task.
__ Transnumeration is not needed to answer this task.

Briefly explain (if necessary):

2. What SKT category does this task address the most? If you believe it addresses multiple categories, check multiple boxes AND explain why you checked each category.

__ Common Knowledge of Content　　__ Specialized Content Knowledge

__ Knowledge of Content and Students　　__ Knowledge of Content and Teaching

Briefly Explain (if multiple boxes checked):

3. To what extent do you think discussing this task with preservice teachers will provide information on their knowledge and use of transnumeration?

__ A Great Deal　　__ Quite a Bit　　__ Somewhat　　__Not at All

Briefly explain (if necessary):

4. Which research questions do you think might be answered, at least partially, by having preservice teachers complete this task? Check all that apply.

    __ 1a) What statistical subject matter knowledge do preservice teachers display with graphical representations?
    __1b) What statistical subject matter knowledge do preservice teachers display with transnumeration between graphical representations?
    __ 2a) How do preservice teachers suggest they will use transnumeration as an inservice teacher?
    __ 2b) What will preservice teachers suggest students know about graphical representations? Transnumeration?
    __ 2c) What pedagogical tactics will preservice teachers suggest to teach concepts in graphical representations?
    __ 2d) What pedagogical tactics do preservice teachers suggest they should use to support students on high-stakes assessments such as an AP exam or an end of course assessment?

5. Please provide specific feedback on revisions for this task including word-choice clarifications (e.g. "I did not understand this phrase . . ."), assessment limitations (e.g. "Your correct answer was also the answer with the most words. Good test taking skills led me to the answer, not knowledge of transnumeration"), math/stats education research (e.g. "Professor Marley did something similar in his research, you should look into that as a source for this question"), or other useful considerations in task development.

Appendix B: Survey Background Questionnaire

Name: ________________________ Current Year at Indiana University: __________

Approximate College GPA: ________ Approximate HS GPA: ___________

Did you take AP-Statistics in High School? ______ (yes/no) if yes, Dual-Credit? _____

if yes, what was your grade: _______ AP-Statistics exam score: _______ / 5

Please write "N" for need to take this course still, "E" for enrolled now, or "C" for completed:

M302 – Mathematics throughout the Secondary Curriculum - Algebra Across the Curriculum (1 Credit)	
M302 - Mathematics throughout the Secondary Curriculum - Probability Across the Curriculum (1 Credit)	
M302 - Mathematics throughout the Secondary Curriculum - Calculus Across the Curriculum (1 Credit)	
M321 - Teaching Mathematics in the Middle School	
M422 – Teaching Mathematics in the Secondary School	

Rate the follow reasons importance to you on why you are becoming a math teacher. Please using 1-5 rankings only 1 time (1 not important, 3 somewhat important, 5 very important):

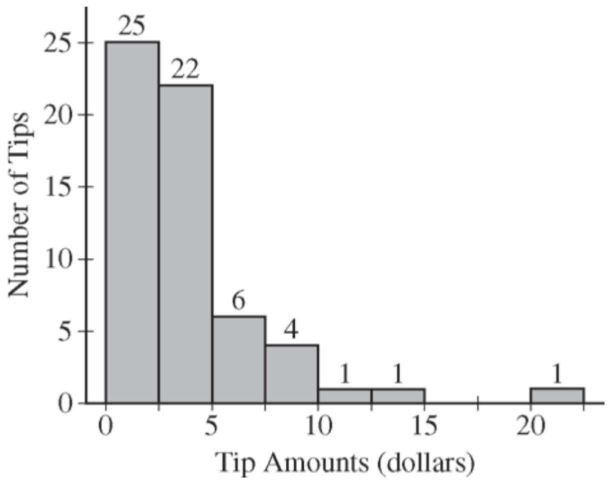
With the math teacher shortage, I believe I will be able to get and keep a job.	1	2	3	4	5
I believe I will be able to control how I get to teach students.	1	2	3	4	5
I believe I have the skills needed to teach mathematics well.	1	2	3	4	5
I believe that my salary and hours of work will give me flexibility in life.	1	2	3	4	5
I believe I can teach mathematics better than statistics.	1	2	3	4	5
I believe my coursework is training me well to become an AP-Statistics teacher.	1	2	3	4	5
I believe I am teaching math for a greater purpose.	1	2	3	4	5
I want to share my love of mathematics with secondary school students?	1	2	3	4	5
List other reasons you choose to become a mathematics teacher if different from above:	________________________________ ________________________________ ________________________________				

Please continue to the following page →

Appendix C: The Tips Task with reflection and belief sections

Johnny is a student of yours who is working on the following problem:

"The histogram below shows a waitress's 60 tip amounts from this past week. One of the tip amounts was mistakenly added to the graph as being $8, when in fact the tip was actually $18. What effect would this mistake have on the mean and median?"



Johnny responds to the question by saying, "The mean and median tip amounts should both increase." Is Johnny correct? What would your next action be in response to Johnny's answer?

A. Johnny is correct. Next, I would ask Johnny to describe the appearance of the distribution because it's important to understanding the data set.
B. Johnny is correct. Next, I would ask Johnny if he thought it would be helpful to track tip amounts for another week to see similarities and differences.
C. Johnny is incorrect. Next, I would have Johnny sum the differences between each tip and the mean and compare them to each other.
D. Johnny is incorrect. Next, I would ask Johnny what effect removing the $20+ tip amount would have on the mean.
E. Johnny is incorrect. Next, I would have Johnny cut each histogram bar out and tape the bars together in order by length. Then have Johnny fold the entire length in half to show the median and discuss how this would change given the clerical mistake.

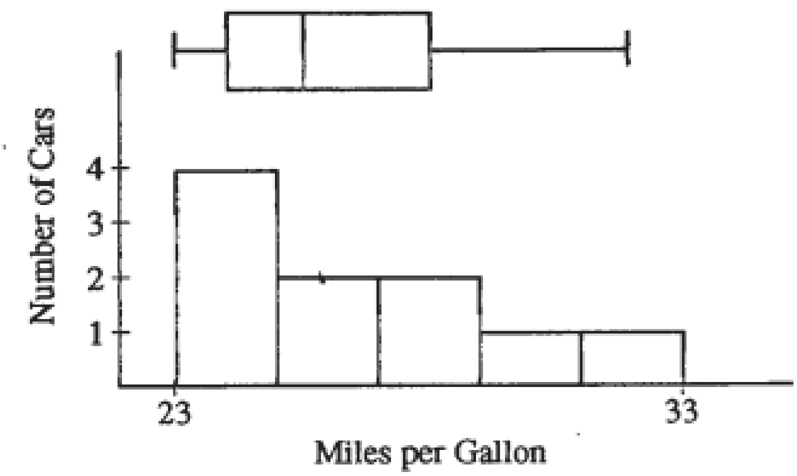
Reflection Section – Please record notes to why you chose the answer(s) above.

Task Belief Questions							
Listed below are several statements about the teaching question you just answered. Please indicate your agreement with each statement by checking the appropriate bubble.							
	Strongly Agree	Agree	Somewhat Agree	Neutral	Somewhat Disagree	Disagree	Strongly Disagree
This question is about a topic I have studied in a college class.	○	○	○	○	○	○	○
I am good at answering questions like this one.	○	○	○	○	○	○	○
I often feel nervous when I try to answer questions like this one.	○	○	○	○	○	○	○
If I try hard, I can usually figure out questions like this one.	○	○	○	○	○	○	○
Secondary mathematics teachers should know how to answer this question.	○	○	○	○	○	○	○

Appendix D: The Fuel Task

Joe is a student of yours that is working on a problem with the following information:

"A consumer organization was concerned that an automobile manufacturer was providing misleading information about average fuel efficiency by saying a new model of car gets 27 miles per gallon. The organization's researchers selected a random sample of 10 cars and assigned each to a random driver for 5,000 miles. The total fuel consumption for 5,000 miles was used to compute the mpg for each car. Below is a boxplot, histogram, and table that records the 10 sample values."



Minimum	Q1	Median	Q3	Maximum
23	24	25.5	28	32

Which of the following statement(s) could Joe make that shows some knowledge about skewness?

A. One way to describe the skewness of the data is the ratio $\frac{sample\ mean}{sample\ median}$. If the population is skewed to the right like above, the mean will be greater than the median, resulting in a large $\frac{sample\ mean}{sample\ median}$ ratio.

B. One way to describe the skewness of the data is the ratio $\frac{sample\ mean}{sample\ median}$. The closer the $\frac{sample\ mean}{sample\ median}$ ratio is to the value of "1," the closer the sample is to being symmetrical.

C. The mean is the vertical line inside the box in the boxplot pictured above.

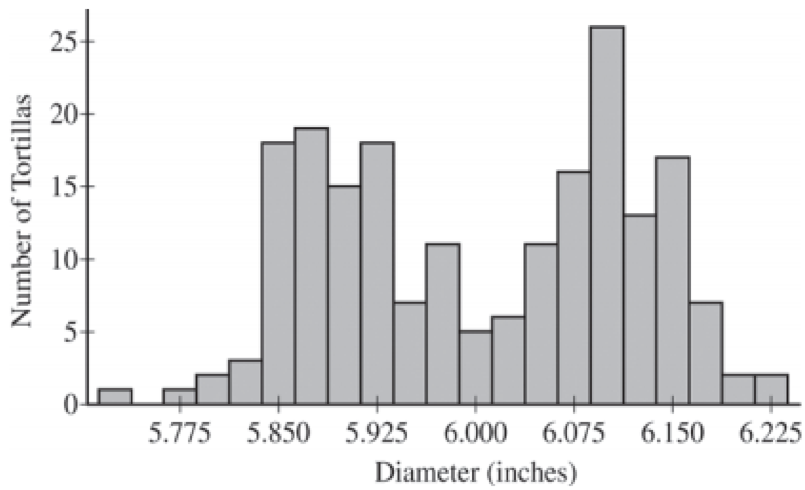
D. Using the table, you could create a formula that describes skewness of the data. A formula that would measure skewness is $\frac{maximum}{Q3}$.

E. Using the table, you could create a formula that describes skewness of the data. A formula that would measure skewness is $\frac{maximum - minimum}{median}$.

Appendix E: The Factory Task

Peggy is a high school student in your class working on the following problem:

A factory's goal is to produce tortillas at 6 inches in diameter. The quality control workers in the factory take a random sample of 200 tortillas across multiple production lines. Results show a sample mean of 6.02 with the following distribution:



Peggy describes the factory's performance saying "The factory is doing a great job of meeting their goal. The sample mean is only .02 inches away from the population mean goal." What should you do next as Peggy's teacher to further her understanding of the factor's performance?

A. The factory's performance is outstanding being only .02 inches away from the mean. As her teacher, I would agree with Peggy and ask if she thinks the company should take another random sample of 200 tortillas to verify results.
B. As Peggy's teacher, I would try to highlight why a histogram is used to represent data. Have Peggy create a Stem and Leaf plot by hand to see how the histogram above was developed.
C. Calculate the mean and median of the data set, and then ask Peggy to compare how the two measures relate to each other. Finish by highlighting the critical differences between the mean and median.
D. Have Peggy draw a normally distributed histogram and says it's mean is also only .02 inches away from 6 inches. Then ask which production line in the factory has better performance based on the two distributions.
E. Peggy should change the scale on the y-axis from an interval of 5 to 1 to get a more precise histogram of the amount of tortillas with each diameter range. Then ask Peggy to explain if the sample mean changed at all given the more detailed histogram.

References

Abrams, S. R., Oppenheim, D. V., Pazel, D. P., & Wright, J. L. (2001). *U.S. Patent No. 6,285,367*. Washington, DC: U.S. Patent and Trademark Office.

American Statistical Association (2017). What's going on with this graph. *The New York Times.* Retrieved from https://www.nytimes.com/column/whats-going-on-in-this-graph.

Agus, M., Penna, M. P., Peró-Cebollero, M., Guàrdia-Olmos, J., & Pessa, E. (2014). The application of graphical representations in estimation of probabilistic events. *Ricerche di Pedagogia e Didattica, Journal of Research and Theories in Education*, *9*(1), 235-252.

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*(2), 131-152.

Andrews, P. (2011). The cultural location of teachers' mathematical knowledge: Another hidden variable in mathematics education research? In T. Rowland & K. Ruthven (Eds.), *Mathematical Knowledge in Teaching* (pp. 99-118). Netherlands: Springer.

Arteaga, P., Batanero, C., Contreras, J. M., & Cañadas, G. (2012). Understanding statistical graphs: A research survey. *Boletín de Estadística e Investigación Operativa*, *28*, 261-277.

Bakker, A., Biehler, R., & Konold, C. (2004) Should young students learn about box plots? In G. Burrill & M. Camden (Eds.), *Curricular Development in Statistics Education: The International Association for Statistical Education 2004 Roundtable* (pp. 163-173*).* Voorburg, Netherlands: International Statistical Institute.

Bakker, A., & Gravemeijer, K. (2005). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147-168). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ball, D. L., Hill, C. H., & Bass, H. (2005). Knowing Mathematics for Teaching: Who Knows Mathematics Well Enough To Teach Third Grade, and How Can We Decide? *American Educator*, (29) 14-22.

Ball, D.L., & Bass, H. (2000). Making believe: The collective construction of public mathematical knowledge in the elementary 155 classroom. In D. Phillips (Ed.), *Yearbook of the National Society for the Study of Education, Constructivism in Education*. 193-224. Chicago, IL: University of Chicago Press.

Batanero, C., Godino, J., & Roa, R. (2004). Training Teachers To Teach Probability. *Journal of Statistics Education* 12(1), 1-19.

Begg, A., & Edwards, R. (1999, December). Teachers' ideas about teaching statistics. Paper presented at the *Combined Annual Meeting of the Australian Association for Research in Education and the New Zealand Association for Research in Education.* Melbourne, Australia.

Ben-Zvi, D., & Garfield, J. B. (Eds.). (2005). *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 3-16). Dordrecht: Kluwer academic publishers.

Boaler, J. (2008). *What's Math Got To Do With It? Helping Children Learn To Love Their Least Favorite Subject--and Why It's Important For America*. Baltimore: Viking/Penguin.

Boaler, J., Dweck, C. S. (2016) *Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages and innovative teaching.* San Francisco: Jossey-Bass.

Bright, G. W., & Friel, S. N. (1998). Graphical representations: Helping students interpret data. *Reflections on statistics: Learning, teaching, and assessment in grades K–12*, 63-88.

Burrill, G. (2014). Preparing future teachers to teach statistics. In K. Makar, B. de Sousa, & R.

Gould (Eds.), *Proceedings of the Ninth International Conference on Teaching Statistics* Flagstaff, Arizona. Retrieved from https://iaseweb.org/Conference_Proceedings.php

Burgess, T. A. (2006). A framework for examining teacher knowledge as used in action while teaching statistics. In Rossman A. & Chance C. (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics, Salvador, Bahia, Brazil*, Retrieved from https://iase-web.org/Conference_Proceedings.php

Burgess, T. A. (2009). Statistical knowledge for teaching: exploring it in the classroom. *For the Learning of Mathematics*, *29*(3), 18-21.

Burgess, T. A. (2011). Teacher knowledge of and for statistical investigations. In Batanero C., Burrill G., & Reading C. (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 259-270). Dordrecht, Netherlands: Springer.

Casey, S. A. (2010). Subject matter knowledge for teaching statistical association. *Statistics Education Research Journal, 9(2), 50-68.*

Cashin, S. E., & Elmore, P. B. (2005). The Survey of Attitudes Toward Statistics scale: A construct validity study. *Educational and Psychological Measurement*, *65*, *509-524*. https://doi.org/10.1177/0013164404272488

Chance, B., Wong, J., & Tintle, N., (2017) Student Performance in Curricula Centered on Simulation-Based Inference: A Preliminary Report. *Journal of Statistics Education. 24(3), 114-126. https://doi.org/10.1080/10691898.2016.1223529*

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, *104,* 801-823. DOI: 10.2307/2975286

College Board. (2016). “AP Statistics Course Description.” *The College Board* [online], Retrieved from https://apstudent.collegeboard.org/apcourse/ap-statistics.

College Board. (2015). “Student Performance Q&A: 2015 AP Statistics Free-Response Questions.” *The College Board* [online], Retrieved from http://media.collegeboard.com/digitalServices/pdf/ap/ap15_statistics_student_performance_qa.pdf.

Common Core State Standards Initiative. (2010). *Common Core State Standards for Mathematics*. Retrieved from http://www.corestandards.org/

Conference Board of the Mathematical Sciences. (2001). *The Mathematical Education of Teachers*. Providence, RI, and Washington, DC: American Mathematical Society and Mathematical Association of America. Retrieved from *https://www.cbmsweb.org/archive/MET_Document/index.htm*

Conference Board of the Mathematical Sciences. (2012). *The Mathematical Education of Teachers II*. Providence, RI, and Washington, DC: American Mathematical Society and Mathematical Association of America. Retrieved from *https://www.cbmsweb.org/archive/MET2/index.htm*

Contreras, J. M., Batanero, C., Díaz, C., & Fernandes, J. A. (2011). Prospective teachers' common and specialized knowledge in a probability task. In D. Pratt, (Ed.), *Proceedings for the Seventh Congress of the European Society for Research in Mathematics Education*, Rzeszow, Poland. Retrieved from http://www.cerme7.univ.rzeszow.pl/index.php

Cooper, L. L., & Shore, S. F. (2010). The effects of data and graph type on concepts and visualizations of variability. *Journal of Statistics Education, 18(2), 1-16*.

Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory, (4th ed.)*. Thousand Oaks, CA: Sage.

Curcio, F. (1987). Comprehension of Mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, *18*, 382-393. doi:10.2307/749086.

Curcio, F. R. (2010). *Developing data graph comprehension (3rd ed.)*. Reston, VA: National Council of Teachers of Mathematics.

delMas, R., Garfield, J., & Ooms, A. (2005). Using assessment to study students' difficulty in reading and interpreting graphical representations of distributions. In K. Makar (Ed.), *Proceedings of the Fourth International Research Forum on Statistical Reasoning, Literacy, and Reasoning* (on CD). Auckland, New Zealand: University of Auckland.

Dweck, C. S. (2008) *Mindset: the new psychology of success.* New York: Ballantine Books.

Edwards T. G., Özgün-Koca A., & Barr, J. (2017). Interpretations of boxplots: helping middle school students to think outside the box. *Journal of Statistics Education 25(1), 21-28.* https://doi.org/10.1080/10691898.2017.1288556

Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal*, *16(1), 44-49.*

Estrada, A., & Batanero, C. (2008). Teachers' attitudes toward statistics. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 163-174). Dordrecht, Netherlands: Springer.

Franklin, C., Bargagliotti, A., Case, C., Kader, G., Scheaffer, R., & Spangler, D. (2015). *Statistical education of teachers.* Alexandria, VA: American Statistical Association. Retrieved from http://www.amstat.org/education/SET/SET.pdf.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report*.

Alexandria VA: American Statistical Association. Retrieved from http://www.amstat.org/education/gaise/

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

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32, 124-158.

Garfield, J., & Franklin, C. (2011). Assessment of learning, for learning, and as learning in statistics education. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 133-145). Dordrecht, Netherlands: Springer.

Gal, I., & Garfield, J. B. (Eds.) (1997). *The assessment challenge in statistics education.* Amsterdam, Netherlands: IOS Press.

Godino, J. D., Ortiz, J. J., Roa, R., & Wilhelmi, M. R. (2011). Models for statistical pedagogical knowledge. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 271-282). Dordrecht, Netherlands: Springer.

González, O. (2016). A framework for assessing statistical knowledge for teaching based on the identification of conceptions of variability held by teachers. In D. Ben-Zvi & K. Makar (Eds.) *The Teaching and Learning of Statistics* (pp. 315-325). Springer International Publishing.

González, M. T., Espinel, M. C., & Ainley, J. (2011). Teachers' graphical competence. In C. Batanero, G. Burrill G & C. Reading (Eds.), *Teaching statistics in school mathematics:*

Challenges for teaching and teacher education (pp. 187-197). Dordrecht, Netherlands: Springer.

González, M. T., & Pinto, J. (2008). Conceptions of four pre-service teachers on graphical representation. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Proceedings of the Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference. Monterrey, Mexico.* Retrieved from https://iase-web.org/

Groth, R. E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education*, *38,* 427-437.

Groth, R. E. (2010). Situating qualitative modes of inquiry within the discipline of statistics education research. *Statistics Education Research Journal*, *9*(2), 7-21.

Groth, R. E. (2012). The Role of Writing Prompts in a Statistical Knowledge for Teaching Course. *The Mathematics Teacher Educator*, *1*(1), 23-40.

Groth, R. E. (2014). Using Work Samples from the National Assessment of Educational Progress (NAEP) to Design Tasks that Assess Statistical Knowledge for Teaching. *Journal of Statistics Education, 22(3)*, 1-28.

Groth, R. E., & Bergner, J. A. (2013). Mapping the structure of knowledge for teaching nominal categorical data analysis. *Educational Studies in Mathematics*, *83*, 247-265.

Haines, B. (2015). Conceptualizing a framework for Advanced Placement Statistics teaching knowledge. *Journal of Statistics Education, 23(3),* 1-16.

Hill, H. C., & Ball, D. L. (2004). Learning Mathematics for Teaching Results from California's Mathematics Professional Development Institutes. *Journal for Research in Mathematics Education, 35*, 330-351.

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

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, *105(1),* 11-30.

Hill, H. C., Sleep, L., Lewis, J. M., & Ball, D. L. (2007). Assessing teachers' mathematical knowledge: What knowledge matters and what evidence counts? In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*, (pp. 111-155). Charlotte, NC: Information Age Publishing.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42,* 371-406.

Jacobbe, T. (2007). *Elementary school teachers' understanding of essential topics in statistics and the influence of assessment instruments and a reform curriculum upon their understanding* (Doctoral dissertation). Retrieved from https://iase-web.org/Publications.

Jacobbe, T., (2011). Elementary School Teachers' Understanding of the Mean and Median. *International Journal of Science and Mathematics Education*, *10,* 1143-1161.

Jacobson, E. J., Audeniz, F., Creager, M. A., Daiga, M., & Uzan, E. (2018) Mathematics Teachers' Knowledge and Productive Disposition for Teaching: A Framework and Measure. In Stylianides G., & Hino K., (Eds.) *Research Advances in the Mathematical Education of Pre-service Elementary Teachers: ICME-13 Monographs* (pp. 187-203). Springer.

Kvale, S. (1996). The 1,000-page question. *Qualitative Inquiry*, *2,* 275-284.

Lancaster, S. M. (2007). *Pre-service teachers and statistics: Interrelationships between content confidence, knowledge, and attitudes, Pedagogical content knowledge, and teacher interest in professional development in statistics*. (Doctoral dissertation). Retrieved from https://iase-web.org/Publications.

Lee, H., Driskell, S. O., Harper, S. R., Leatham, K. R., Kersaint, G., & Angotti, R. L. (2011). Prospective Teachers' Use of Representations in Solving Statistical Tasks with Dynamic Statistical Software. In L. Wiest & T. Lamberg (Eds.) *Proceedings of the 33rd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Reno, Nevada.

Lee, H., Kersaint, G., Harper, S., Driskell, S., Jones, D., Leatham, K., Agnotti, R., & Adu-Gyamfi, K., (2014). Teachers' use of Transnumeration in solving statistical tasks with dynamic statistical software. *Statistics Education Research Journal*, *13*(1), 25-52.

Lem, S., Onghana, P., Verschaffel, L., & Van Dooren, W. (2013). External representations for data distributions: In search of cognitive fit. *Statistics Education Research Journal 12*(1) 4-19.

Kaplan, J., Gabrosek, J., Curtiss, P., & Malone, C. (2014) Investigating Student Understanding of Histograms. *Journal of Statistics Education*. *22*(2) 1-30.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis* (second edition). Beverly Hills, CA: Sage.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: US Department of Education.

Nasser, F. M. (2004). Structural model of the effects of cognitive and affective factors on the achievement of Arabic-speaking pre-service teachers in introductory statistics. *Journal of Statistics Education*, *12*(1), 1-28.

Noll, J. (2007). *Statistics teaching assistants' statistical content knowledge and pedagogical content knowledge.* Unpublished doctoral dissertation. Portland, OR: Portland State University.

Norman, G., (2010) Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education, 15*(5), 625-632.

Olfos, R., Estrella, S., & Morales, S. (2014). Open Lessons Impact Statistics Teaching Teachers' Beliefs. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Proceedings of the Ninth International Conference on Teaching Statistics, Flagstaff, Arizona, USA.* Retrieved from iase-web.org.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press.

Philipp, R. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257–318). Charlotte, NC: Information Age Publishing.

Pierce, R. & Chick, H., (2011) Teachers' Beliefs About Statistics Education. In C. Batanero,

G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics: Challenges for teaching and teacher education* (pp. 151-162). Dordrecht, Netherlands: Springer.

Roberts, D. M., & Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement*, *40*, 235-238.

Rodway, P., Schepman, A., & Thoma, V. (2016) Reachability does not explain the middle preference: A comment on Bar-Hillel (2015), *i-Perception*, *7*(2), 1-5.

Rossman, A., Chance, B., & Medina, E. (2006). Some important comparisons between statistics and mathematics, and why teachers should care. In G. F. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance: Sixty-eighth annual yearbook of the National Council of Teachers of Mathematics* (pp. 323-333). Reston, VA: NCTM.

Roth, W. M., & Bowen, G. M. (2001). Professionals read graphs: A semiotic analysis. *Journal for research in mathematics education*, *32,* 159-194.

Rouan, O. (2002). Secondary School Math Teachers' conceptions of the statistical graphics functions, reading and interpretation. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town, South Africa.

Scheaffer, R. L. (2006). Statistics and Mathematics: On Making a Happy Marriage. In G. F. Burrill & P. C. Elliot (Eds.), *Thinking and reasoning with data and chance: Sixty-eighth annual yearbook of the National Council of Teachers of Mathematics* (pp. 309-321). Reston, VA: NCTM.

Shaughnessy, J. M. (1992). Research in probability and statistics: reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: Macmillan.

Shaughnessy, J. M. (2007). Research on Statistics Learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning.* (pp. 957-1009). Charlotte, NC: Information Age.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. Bishop (Ed.), *International handbook of mathematics education* (pp. 205-237). Dordrecht, The Netherlands: Kluwer.

Shaughnessy, J. M., & Pfannkuch, M. (2002). How faithful is Old Faithful? Statistical thinking: A story of variation and prediction. *Mathematics Teacher*, *95*(4), 252.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4-14.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.

Turner, F., & Rowland, T. (2011). The Knowledge Quartet as an organizing framework for developing and deepening teachers' mathematics knowledge. In T. Rowland & K. Ruthven (Eds.), *Mathematical knowledge in teaching* (pp. 195-212). New York: Springer.

Watson, J. M. (2001). Profiling teachers' competence and confidence to teach particular mathematics topics: The case of chance and data. *Journal of Mathematics Teacher Education*, *4*(4), 305-337.

Watson, J. M. (2013). *Statistical literacy at school: Growth and goals*. Mahwah, NJ: Lawrence Erlbaum Associates.

Watson, J. M., & Moritz, J. B., (1998). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics, 37*, 145-168.

Watson, J., Callingham, R. A., & Donne, J. M. (2008). Establishing pedagogical content knowledge for teaching statistics. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.) *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.* Monterrey, Mexico. Retrieved from https://iase-web.org/

Watson, J. M., Callingham, R., & Nathan, E. (2009). Probing teachers' pedagogical content knowledge in statistics: How will Tom get to school tomorrow? In R. Hunter, B. Bicknell, & T. Burgess (Eds.), *Proceedings of the 32nd Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 563-570). Palmerston North, New Zealand.

Wessels, H., (2014). *Developing statistical knowledge for teaching of variability through professional development.* In K. Makar, B. de Sousa, & R. Gould (Eds.), *Proceedings of the Ninth International Conference on Teaching Statistics, Flagstaff, Arizona, USA.* Retrieved from iase-web.org

Whitaker D., & Jacobbe, T. (2017) Students' understanding of bar graphs and histograms: Results from the LOCUS Assessments. *Journal of Statistics Education, 25(2)*, 90-102.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*, 223-24.