STUDENTS' PERCEPTIONS OF STATISTICS: AN EXPLORATION OF ATTITUDES, CONCEPTUALIZATIONS, AND CONTENT KNOWLEDGE OF STATISTICS

MARJORIE E. BOND
Monmouth College
mebond@monmouthcollege.edu

SUSAN N. PERKINS
Northwest Nazarene University
sperkins@nnu.edu

CAROLINE RAMIREZ
University of the Pacific
caaramirez@aol.com

ABSTRACT

Although statistics education research has focused on students' learning and conceptual understanding of statistics, researchers have only recently begun investigating students' perceptions of statistics. The term perception describes the overlap between cognitive and non-cognitive factors. In this mixed-methods study, undergraduate students provided their perceptions of statistics and completed the Survey of Students' Attitudes Toward Statistics-36 *(SATS-36). The qualitative data suggest students had basic knowledge of what the word statistics meant, but with varying depths of understanding and conceptualization of statistics. Quantitative analysis also examined the relationship between students' perceptions of statistics and attitudes toward statistics. We found no significant difference in mean pre- or post-SATS scores across conceptualization and content knowledge categories. The implications of these findings for education and research are discussed.*

Keywords: *Statistics education research; SATS-36; Student attitudes; Conception of statistics*

1. INTRODUCTION

Most undergraduate majors require a statistics course as a pre-requisite or in fulfilment of general education requirements. An undergraduate statistics course provides important resources for functioning effectively in environments that value information and numeracy because these are central in making informed decisions based on numerical data (Gal, 2002; Utts, 2003). As the field of statistics education research is building, it would be helpful to know whether students preparing to take an introductory statistics course at the undergraduate level know what is meant by the term "statistics," and how that conceptualization differs at the end of the course.

Existing research does not tell us enough about student perceptions of statistics, and research findings may be confounded with students' beliefs and attitudes toward mathematics (Gal, Ginsburg, & Schau, 1997). Many of the researchers in statistics education collect their survey data from students who have just enrolled in an introductory statistics course, which may affect how students interpret the word statistics, as pointed out by Gal et al.:

> Since almost all of the items on most attitude surveys include the word "statistics," it is important to realize that some high school or would-be college students convey some fuzziness regarding what the term "statistics" might be about or about life domains where statistics may be used. How this "fuzziness" affects the validity or usefulness of surveys of precollege students is thus a matter for some concern. (p. 6)

Statistics Education Research Journal, 11(2), 6-25, http://www.stat.auckland.ac.nz/serj

For researchers in statistics education, this presents an opportunity to assess how students regard the term "statistics" at the beginning of the semester and toward the end. A better understanding of students' definitions of the term "statistics" could also extend the discussion of the validity and usefulness of surveys that include the word "statistics." In this article, we use the term *perception* to describe the overlap between cognitive (defining and/or conceptualizing statistics) and non-cognitive (attitude or motivation) factors. This study was designed to answer the following research questions:

1. How do students define and conceptualize statistics at the beginning of the semester in an elementary statistics course?
2. Are there changes in student definitions and conceptualizations from the beginning of the course to the end of the course? If so, to what extent are these different?
3. What is the relationship between student definition of statistics and attitudes toward statistics?
4. What is the relationship between student conceptualization of statistics and attitudes toward statistics?

2. LITERATURE REVIEW

We first discuss how the statistics education community has defined statistics. Second, we review current research in how students define and conceptualize statistics. We end this section with a review of the literature on students' attitudes in statistics courses.

2.1. WHAT IS STATISTICS?

How do researchers in statistics education and statisticians define statistics? Although statistics is often viewed as a branch of mathematics, statistics is a discipline that involves more non-mathematical activities than the actual use of mathematics (Cobb & Moore, 1997; DeVeaux & Velleman, 2008; Higgins, 1999). As DeVeaux and Velleman noted, the challenge in teaching statistics is that "we have a wide variety of skills to teach, and most of them require judgment in addition to mathematical manipulation" (p. 55). Furthermore, we live in a society where information and numerical data have played an increasing role in matters of policy and decision making. As a result, our mathematics and statistics education communities have a civic responsibility to develop our students' statistical literacy, statistical reasoning, and statistical thinking (Gal, 2002; Utts, 2003; Wallman, 1993).

In Wallman's (1993) presidential address at the 1992 annual meeting of the American Statistical Association, she defined statistical literacy as the "ability to understand and critically evaluate statistical results that permeate our daily lives – coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions" (p. 1). From this definition, statistical literacy involves an *appreciation* of statistics, which can only come from a person's psychological mindset or disposition. Gal (2002) identified two dispositional elements important in developing statistical literacy; these are: (1) beliefs and attitudes, and (2) critical stance. Similarly, Watson (1997) argued that statistical literacy involves a three-tiered hierarchy of skills: basic understanding of statistics, an understanding of statistics based on the context, and a questioning attitude. Watson's model of statistical thinking emphasized both the cognitive (understanding) and affective (attitudes) factors needed for students to develop statistical literacy (Watson & Callingham, 2003). The latest iteration of Watson's framework of statistical literacy included a hierarchy of six constructs: (1) idiosyncratic, (2) informal, (3) consistent, (4) non-critical, (5) critical, and (6) critical mathematical (Watson & Callingham, 2003). This hierarchy was similar to Gal's list of five knowledge bases required in developing statistical literacy, namely: (1) literacy skills, (2) statistical knowledge, (3) mathematical knowledge, (4) context knowledge, and (5) knowledge in posing critical questions.

Whereas statistical literacy relies on both cognitive and non-cognitive factors, statistical reasoning, in contrast, focuses more on cognitive processes, as reflected in a person's active engagement with the data such as interpreting graphs and summary statistics (Ben-Zvi & Garfield, 2004). Garfield and Gal (1999) defined statistical reasoning as "the way people reason with statistical ideas and make sense of statistical information" (p. 207). This definition suggested that statistical reasoning and literacy were separate but not orthogonal (or independent); that is, one would need to be statistically literate in order to show statistical reasoning.

Ben-Zvi and Garfield (2004) defined statistical thinking as "understanding of why and how statistical investigations are conducted… when and how to use appropriate methods of data analysis such as numerical summaries and visual displays of data" (p. 7). In conducting statistical investigations, higher levels of thinking are needed in order to understand and use the context in the data analysis, especially in the interpretation of the statistical results. Thus, statistical thinking requires both statistical literacy and statistical reasoning.

Reform efforts in undergraduate statistics courses, spearheaded by the Mathematical Association of America and the American Statistical Association, recognized the importance of fostering statistical literacy, reasoning, and thinking. Several statistics educators have encouraged others to focus on statistical thinking, reasoning, context, and concepts such as variability rather than mathematics and formulas (Cobb, 1992; Higgins, 1999; Moore, 1997; Rossman, Chance, & Medina, 2006). Statistics' dependence on data and context led statisticians to assert that statistics is a separate discipline from mathematics (Cobb & Moore, 1997; delMas, 2004; DeVeaux & Velleman, 2008). Given this distinction, Cobb and Moore noted the usefulness of statistics in other disciplines and offered a brief definition:

> Statistics is a methodological discipline. It exists not for itself but rather to offer to other fields of study a coherent set of ideas and tools for dealing with data. The need for such a discipline arises from *the omnipresence of variability*. … Statistics provides means for dealing with data that take into account the omnipresence of variability. (p. 801)

However, not all students view statistics with the concept of variability in mind. In fact, some students find it difficult to acknowledge and describe spread or variation in a sample (Reading & Shaughnessy, 2004). The notions of statistics as a "mathematical science" and as the study of variability assisted us in clarifying a focus for understanding the definition of statistics, along with statistical literacy, reasoning, and thinking. Because the above definitions were more technical and were developed by statisticians and statistics educators, our goal with this study was to give voice to students' definition of statistics, and to examine students' perceptions of statistics.

How do undergraduate students in statistics courses define statistics? Only a few researchers have looked at students' perceptions and definitions of statistics. These studies included two qualitative studies: one study collected data via interviews with a small number of students (Reid & Petocz, 2002), and another collected brief, written responses from a larger number of participants (Gordon, 2004). Reid and Petocz interviewed 20 first-year and third-year students taking elementary statistics and regression analysis respectively. Their study resulted in the following six categories of how students defined statistics, organized into three major themes:

By focusing on techniques: (Gathering – Extrinsic Technical)
(1) Statistics is individual numerical activities
(2) Statistics is using individual statistical techniques
(3) Statistics is a collection of statistical techniques
By focusing on data: (Applying – Extrinsic Meaning)
(4) Statistics is the analysis and interpretation of data
(5) Statistics is a way of understanding real-life using different statistical models.
By focusing on meaning: (Creating – Intrinsic Meaning)
(6) Statistics is an inclusive tool used to make sense of the world and develop personal meanings.

Similarly, Gordon (2004) developed five categories to describe how 250 psychology students defined statistics. These were: (1) no meaning, (2) process or algorithms, (3) mastery of statistical concepts and methods, (4) tool for getting results in real life, and (5) critical thinking. Gordon also found that most students had a negative view of statistics.

In both of these previous studies, the participants were already enrolled in an undergraduate statistics course, which likely influenced their definition of statistics. Although providing valuable information, both studies limited their understanding of student perceptions of statistics by only considering the cognitive component and not inquiring about student attitudes.

2.2. ATTITUDES OF STUDENTS TOWARD STATISTICS

Thus far, we have argued that statistical literacy involves cognitive and affective factors. Reform efforts in statistics education emphasized the need to develop students' statistical literacy, reasoning, and thinking. Furthermore, Gal et al. (1997) noted that as more alternative assessment strategies and reform teaching methods are used in the classroom, more research is needed to understand students' attitudes, beliefs, and motivation because these non-traditional learning contexts are more likely to cause affective responses than the more familiar traditional curricula.

Researchers have used a variety of approaches to assess students' attitudes toward statistics. Traditionally, these were self-report measures, such as Likert-scale questionnaires. One of these instruments examined the relationship between *student attitudes* and *conceptions* (Evans, 2007). This instrument, called *Student Attitudes and Conceptions in Statistics* (STACS), required students to interpret or apply conceptual knowledge in evaluating statements on probability and descriptive statistics. Results of this study showed a significant correlation between positive attitudes and accurate conceptions about statistics toward the end of the course. However, we did not use the STACS instrument because it does not address the multidimensional nature of attitudes as it only uses a single score in measuring attitudes. Also, the assessment of conceptions was limited to the students' responses to the Likert-scale items.

To address the limitations, we used the *Survey of Attitudes Toward Statistics* (SATS; Schau, 1992, 2003a), a well-known survey in statistics education, with several authors documenting solid psychometric properties for SATS scores (Dauphinee, Schau, & Stevens, 1997; Hilton, Schau, & Olsen, 2004; Schau, 2003b; Schau, Stevens, Dauphinee, & Del Vecchio, 1995). Extensive work critically analyzing the SATS instrument has supported the reliability, validity, and multi-dimensionality of the scores and constructs (Chiesi & Primi, 2010; Coetzee & Van der Merwe, 2010; Sorge & Schau, 2002; Tempelaar, Gijselaers, Schim van der Loeff, & Nijhuis, 2007; Tempelaar, Schim van der Loeff, & Gijselaers, 2007; Vanhoof, Kuppens, Sotos, Verschaffel, & Onghena, 2011). To a large extent, researchers in statistics education have been using this instrument to assess students' attitudes across various educational settings, interventions, and instructional approaches (Carlson & Winquist, 2011; Carnell, 2008; Dempster & McCorry, 2009; Posner, 2011).

3. METHODOLOGY

Qualitative research methodology provided several possibilities for exploration, including allowing participants to provide their own perceptions, creating space for new ideas, and examining emerging areas of research (Creswell, 1998; Guba & Lincoln, 1981; Kazdin, 1998; Maxwell, 1998). Groth (2010) discussed the importance of using qualitative research methods for multiple purposes in statistics education research and Kalinowski, Lai, Fidler, and Cumming (2010) highlighted the value of mixed methods research. In comparison to qualitative projects which explore completely new areas of research and use extensive data collection, this project allowed us to extend current research by focusing on better understanding an area of statistics education that has already been explored. Because this project had a narrow focus, the qualitative data collected were limited to information needed to answer the research questions.

A recent trend in statistics education research has been the increase of qualitative research, with many studies including mixed methods or multi-method research designs. This trend was noted by delMas (2011) in his keynote address at the United States Conference On Teaching Statistics (USCOTS) in which he provided an overview of the growth and trends in research on statistics education and highlighted the valuable role of qualitative research methodology in research on statistics education. Additionally, in November of 2010, the *Statistics Education Research Journal* published a special issue on using qualitative research methodologies to study statistics education.

3.1. MEASURES

Perception of statistics The first author developed a short-answer survey titled *Perception of Statistics* to collect qualitative and quantitative information on participants' understanding of the term "statistics." This survey had two versions, pre and post, designed to be taken by students prior to and after taking an undergraduate introductory statistics course. Both versions were created on Survey Monkey. Participants typically used 5 to 12 minutes to complete either survey. *Perception of Statistics* was developed from a pilot survey administered to undergraduate statistics students. Based on data analysis from the pilot test and from colleague consultation, the author revised the original survey to enhance clarity. For example, students answered the pilot survey's question of "What do you expect to learn in this course?" with "Statistics." To avoid this uninformative answer, this question was revised to "List 4 to 6 topics which you expect will be discussed in an introductory statistics course. That is, list what you expect to learn." Many questions were reworded to ensure exploration of students' perception of statistics.

After two questions asked about students' past statistics course history, the pre-version of *Perception of Statistics* contained four questions:

(1) What do you think when you hear the word "Statistics?"
(2) List 4 to 6 topics which you expect will be discussed in an introductory statistics course. That is, list what you expect to learn.
(3) How would you define "Statistics?" That is, what is "Statistics?"
(4) What type of work would a person who studied Statistics do? That is, what does a statistician do?

The post version contained six questions:

(1) What do you think when you hear the word "Statistics?"
(2) List 4 to 6 topics which you covered in your introductory statistics course.
(3) Was there any topic(s) covered in this course which you didn't expect? If so, what was it?
(4) Was there any topic(s) that you thought would be covered in this course BUT was not covered? If so, what was it?
(5) How would you define "Statistics?" That is, what is "Statistics?"
(6) What type of work would a person who studied Statistics do? That is, what does a statistician do?

Survey of Attitudes Toward Statistics (SATS-36) The most recent revision of the *Survey of Attitudes Toward Statistics* (SATS-36) contains 36 items which were designed to measure undergraduate students' attitudes toward statistics (Schau, 2003b). There are two versions of the SATS, a pre-course version and a post-course version. The 36 items comprise six subscales, *Affect* (6 items), *Cognitive Competence* (6 items), *Value* (6 items), *Difficulty* (7 items), and the most recent two subscales, *Interest* (4 items) and *Effort* (4 items). The *Affect* subscale measures positive and negative feelings toward statistics. The *Cognitive Competence* subscale measures participants' attitudes regarding their perception of their ability to mentally comprehend statistics. The *Value* subscale measures participants' perceptions of the usefulness and worth of statistics. The *Difficulty* (perceived easiness) subscale is measured by items which collectively asked about participants' attitudes regarding how difficult statistics is/was. The *Interest* subscale measured how much interest a participant has in statistics. Finally, the *Effort* subscale asks about the amount of work participants expect to spend learning statistics. The SATS instruments and scoring guides are available at http://www.evaluationandstatistics.com.

3.2. PARTICIPANTS

Forty-seven participants from a small liberal arts college in the United States completed the pre-course data collection. Twenty-one (44%) were male and 26 (55%) were female. Fifty-one students could have taken the pre-course surveys. Four participants were removed due to failure to obtain complete results on both the pre-perception survey and the pre-SATS-36 survey. The response rate for the pre-surveys was 92%. Ten participants remained in the data set even though they only provided information on the pre-course surveys. Forty-three students could have taken the post-course surveys. Thirty-seven took the post-perception survey and 38 took the post-SATS-36 survey; 16 (42%) were male and 22 (58%) were female. The response rate for the post-course surveys was 86% (perception) and 88% (SATS-36).

Of those who did not complete the post-course data collection, five participants were unavailable for data collection because they withdrew from the course or did not attend class for an extended period of time. Of the four students who completed the course but chose not to participate in the post-course data collection, three received grades of F for the course and the fourth received a C-. Three participants were eliminated due to substantial amounts of missing data. Table 1 reports participants' prior statistics courses.

Table 1. Previous statistics courses

Category	Count	Percentage (n = 47)	Course (if given)
No statistics courses in either High School nor College	36	77%	N/A
Statistics course - High School Only	7	15%	AP Statistics (2); Functions, Statistics, Trigonometry (1)
Statistics course - College Only	3	6%	Retaking this course (2)
Statistics course - High School and College	1	2%	Functions, Statistics, Trigonometry, Basic Statistics

3.3. PROCEDURES FOR COLLECTING DATA

Solicitation Students enrolled in two undergraduate introductory statistics courses in Spring 2011 were invited to participate in this research project. They were informed of this project by their statistics instructor during their first class period. The instructor then provided potential participants with the web links for the pre-course versions of the SATS-36 and for the *Perceptions of Statistics* survey. Participants completed the surveys online within three days of the beginning of the semester. At the end of the semester, students in the course were again invited to participate in the research study and were provided the web links to the post-course versions of the *SATS* and the *Perceptions of Statistics* questionnaire. Students had from the last week of class until the final exam to complete the surveys. They did not know their final grades before completing the surveys. Participation was voluntary; no incentives were offered and there was no penalty for not participating. Prior to beginning the study, permission was obtained from the institution's Institutional Review Board. Students who chose to participate completed the surveys, which took approximately 10 to 20 minutes total for the two surveys.

3.4. ANALYSIS

Content accuracy of statistics definitions To clearly answer the research questions, analysis of the qualitative data initially focused on determining the accuracy of participant responses. In this analysis, we used what Miles and Huberman (1994) define as a *tight data analysis* approach. In tight data analysis, data are analyzed using a pre-existing schema, which can come from previous research studies or theory. We defined each component of the schema and then read the qualitative data to see which classification best fit each piece of data. Data which did not fit the schema were then analyzed separately.

Because the concept of statistics was not explained by previous researchers or statisticians in a simple definition, we determined that the best representation of participants' understanding of statistics would be captured by an ordinal ranking of participants' overall definitions, in keeping with Watson's model of statistical thinking (Watson & Callingham, 2003). In this initial tight data analysis approach, we considered previously published definitions of statistics. Although no one author provided a coding scheme for definitions of statistics, two main themes from previous definitions provided a basis for the coding scheme. These included an emphasis on the value of variability (Cobb & Moore, 1997) and the importance of application and context (Higgins, 1999; Rossman et al., 2006; Wallman, 1993). Additionally, Watson's framework of statistical literacy (Watson & Callingham, 2003) included a hierarchical component. We utilized these themes and the overarching topics in traditional undergraduate statistics courses to read the data and develop the coding scheme for definitions of statistics. We each read the qualitative data completely multiple times and then agreed on a ranking of 1 = probability only or sports statistics, 2 = descriptive statistics only, 3 = descriptive statistics with the concept of variability added, 4 = descriptive statistics with the concept of probability added, and 5 = inferential statistics and/or hypothesis testing (which indicates that the definition includes the application, usefulness, or context of statistics).

Two of the authors read the qualitative data and separately coded participant comments. Participants' comments for all items on the questionnaire were considered collectively when determining the code for each participant's responses. The highest applicable code was given. When later matching the coding scheme to each participant comment, we noticed that some participants' responses overtly stated that the participant did not know what statistics was. We realized that these responses were not captured in our initial coding scheme. Therefore, we added a category which we coded as 0.

On the pre-course data, the two coding authors had an agreement percentage of 55%, with agreement based on both authors assigning the same number code to a participant. After discussing specific data, agreement was 64%. The authors determined that the reason for most of the disagreements was the different professional contexts of the researchers. One of the coders was trained in statistics education and teaches statistics to undergraduates, whereas the other was trained in a human development field and teaches statistics to master's level students. The remaining author discussed coding disagreement. She was chosen for this role because of her experience in the field of statistics and her advanced understanding of statistics concepts. In most cases, reaching an agreement was simple. In every case, the final decision was acceptable to all authors. In discussions, we believed that our diverse perspectives led to more in-depth conversations about statistics education and that our final code definitions were more exact, concise, and thoughtful than they would have been if our initial perspectives had been more congruent and we had initially achieved a higher rate of agreement. The same tight data analysis procedure was used for analyzing the post-course qualitative data, with the two coding authors initially having an agreement percentage of 69%. At that point, categories were clarified via discussion with all three authors. Again, acceptable decisions were reached for each participant.

Student concepts of statistics While reading the qualitative data, the authors agreed that only considering the content accuracy missed much of the richness that participants' responses were providing. Specifically, only examining the accuracy of the definitions did not capture students' conceptualizations of statistics. We noticed that even among participants with the same level of accuracy in definitions, there was variation in what they understood statistics to be and their sense of the usefulness of statistics. We decided to analyze the data again, but to use *loose analysis* (Miles & Huberman, 1994). In loose analysis, we approached the data with as open a mind as possible and allowed the themes to emerge directly from the data. For the tight analysis, the researchers had approached the data looking for definitions. In this loose analysis stage, the researchers approached the data more openly and allowed the participants' comments to go in whatever direction the comments seemed to go. We used the constant comparative method because, as Merriam (1998) noted, it fits well with most genres of qualitative analysis. In using the constant comparative method, we read a section of qualitative data, considered a code that might fit this data, then read another piece of data, considered how that data may or may not fit with previous codes, and adjusted the codes as appropriate throughout the process (Patton, 2005). This process was redundant, iterative, and recursive, with the researchers often coding the same data several times (Merriam, 1998; Patton, 2005; Rossman & Rallis, 2003).

After examining the codes that emerged, we recognized that the codes were very similar to Reid and Petocz's (2002) hierarchical conceptions codes (see Section 2.1). Our codes were 1. math; 2. analysis (manipulation of numbers); 3. application (using numbers to represent something, e.g., sports statistics); 4. research (statistics works with research design principles to be more accurate or useful); 5. meaningfulness (using data to make sense of the world); and 6. impactful (making decisions based on information from research and statistics). Reid and Petocz's (2002) codes included three hierarchical levels, with the first two levels having hierarchical categories. In Reid and Petocz's first (lowest) level, participants approached statistics as a task of gathering information. The first level was "A focus on techniques" (p. 6), with the categories including "1. Statistics is individual numerical activities. … 2. Statistics is using individual statistical techniques" (p. 6), and "3. Statistics is a collection of statistical techniques" (p. 7). In the second (middle) theme, participants applied data. The middle theme was “A focus on using data” (p. 8), and included the categories of “4. Statistics is the analysis and interpretation of data” (p. 8) and “5. Statistics is a way of understanding real-life using different statistical models” (p. 9). In the third (highest) level of “A focus on meaning” (p. 10), participants approached statistics as a way of creating. The singular category in this theme was “6. Statistics is an inclusive tool used to make sense of the world and develop personal meanings” (p. 10).

Based on two considerations, 1. the large similarity between our emerging codes and Reid and Petocz’s (2002) codes, and 2. the value of building on previous literature, we decided to re-code the qualitative data using a tight coding approach. Using tight coding allowed this study to build on the findings of previous research and to determine where current data overlapped with previous research or presented new ideas. In this case, the codes of Reid and Petocz overlapped completely with the data. The process for the tight coding followed the procedures explained above.

Trustworthiness of methods used Kalinowski et al. (2010) assert that qualitative researchers establish trustworthiness by being clear about the researchers’ perspective, providing adequate information about participants, using examples to support the findings, using multiple people for coding to ensure accuracy, writing coherently, focusing on the goals of the research, and connecting with readers. We have attempted to establish trustworthiness by focusing on each of these techniques throughout this article.

Quantitative analysis of data Because the codes used in the qualitative analysis were ordinal and simple, that is, the codes were ordered in levels of increasing accuracy or complexity and did not include themes across categories, the qualitative codes were easily converted into quantitative variables. These variables included (1) *Content*, which described the content and accuracy of participants’ definitions of statistics; and (2) *Concept*, which described participants’ own conceptualizations of statistics.

Participants’ responses on the SATS items were reverse coded when appropriate and averaged to create subscale scores. Subscale scores were then used in quantitative analyses. Quantitative analyses began with descriptive statistics, and the sign test was used to examine differences between pre- and post-course Content and Concept scores. We used Spearman’s rho correlations to examine the relationships between variables because the data were ordinal. Because four correlations were calculated, the Bonferroni correction was used $\alpha = 0.05/4 = 0.0125$ for the individual tests, which kept the alpha level at most 0.05. We used ANOVA to examine whether there were significant differences in means of SATS scores across groups of participants with different Content scores. Similarly, we used ANOVA to test differences in means of SATS scores across groups of participants with different Concept scores. These ANOVA involve 24 tests, so to ensure an overall significance level of at most 0.05 using a Bonferroni correction, we require significance levels of $\alpha = 0.05/24 = 0.002$ for the individual tests.

4. RESULTS

Qualitative results for this study focused on six levels of the content of participants’ definitions of statistics and their conceptualizations of statistics. Quantitative analyses compared pre- and post-course Content and Concept scores, and analyzed the relationship between participants’ reported SATS scores and Content and Concept scores. Finally, participants’ definitions, conceptualizations, and attitudes were considered together to comprise perceptions of statistics. Any potentially identifying information was removed and participants were referred to by randomly-assigned gendered pronouns for ease of writing purposes.

4.1. STUDENT DEFINITIONS AND CONCEPTUALIZATIONS OF STATISTICS

The qualitative analysis answered the first research question: How do students define and conceptualize statistics at the beginning of the semester in an elementary statistics course? Although we initially analyzed the pre-test data separately from the post-test data, the qualitative analysis indicated that students’ definitions for both the pre-test and post-test data contained the same categories. Therefore, we used the same coding scheme for both pre-test and post-test data. Table 2 reports the categories and number of responses in each category. Figures 1 and 2 illustrate the percentages of responses under each category of "Content" and "Concept."

Content of participants' definitions These analyses resulted in six cumulative levels of Content, arranged in order of lowest conceptual understanding of statistics to highest. The first two categories were determined to be inaccurate definitions of statistics; the third category (descriptive statistics only) was considered an accurate, but simplistic definition of statistics, with the final three categories reflecting accurate definitions. Table 2 and Figure 1 show the responses of participants whose comments fit in each category. Table 3 lists the categories coded as "Content," describes each category, and provides quotes to illustrate participant comments which were coded in each category. [Spelling and grammar errors in participants' original responses were corrected for reporting purposes.] Table 3 also lists the quantitative value each category was assigned when Content was converted to a quantitative variable.

Table 2. Report of participants' content knowledge and conceptualizations of statistics

Content Knowledge (Identify topics about statistics)	Pre (*n* = 47)		Post (*n* = 37)	
	Count	Percentage	Count	Percentage
No understanding	4	9%	0	0%
Probability only; sports statistics	7	15%	0	0%
Descriptive statistics only (mean, median, mode)	20	43%	1	3%
Descriptive statistics with emphasis in variability	6	13%	3	8%
Descriptive statistics and probability	5	11%	11	30%
Inferential statistics	5	11%	22	60%
Conceptualizations (Statistics is about…)	**Pre (*n* = 47)**		**Post (*n* = 37)**	
	Count	Percentage	Count	Percentage
TECHNIQUES				
Equations	11	23%	0	0%
Using individual techniques	8	17%	0	0%
Using a collection of techniques	7	15%	7	19%
USEFULNESS				
Analysis and interpretation of data	12	26%	18	49%
Understanding real-life using different statistical models	6	13%	10	27%
MEANINGFULNESS				
Making sense of the world and developing personal meanings	3	6%	2	5%

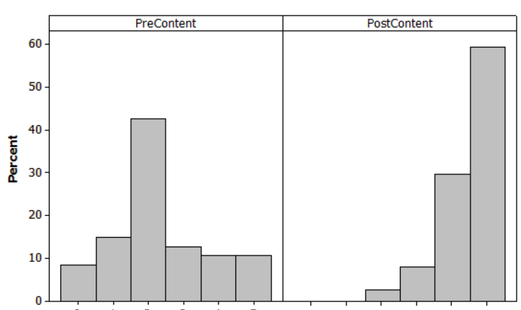

Figure 1. Histograms of Content (Pre and Post)

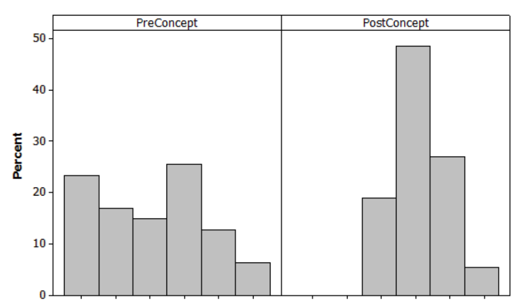

Figure 2. Histograms of Concept (Pre and Post)

Conceptualization in participants' definitions Reid and Petocz' (2002) themes of students' conceptions of statistics had three hierarchical levels: 1. Techniques, 2. Using Data, and 3. Meaningfulness. Each level contained sub-categories, which were also hierarchical. Table 2 summarizes the responses of participants' conceptualizations of statistics in the pre- and post-surveys. The dramatic difference between the pre- and post- responses can be seen in Figure 2. Table 4 lists the category names, descriptions, and illustrative quotes of the categories under Concepts. The table also includes the quantitative values used for each category when Concepts was converted to a quantitative variable.

Table 3. Content categories and quotes

Quantitative Value	Category Name	Description	Illustrative Quote(s)
0	No understanding	The participant did not understand that statistics was separate from math. Inaccurate definition	(a) Honestly, I know nothing except the mathematics. (b) When I hear the word statistics I think about a combination of pre-algebra and geometry and trig. … Statistics is a way of figuring out mathematical equations and solutions.
1	Probability	Participants described some component of probability, with several participants listing sports statistics. Inaccurate definition	I think of the statistics of sports, as in scores and records. … Sports scores, probability ratios, fractions. I consider statistics as ratios or fractions. As in, a percent of one thing, out of a whole.
2	Descriptive	"Statistics" was understood as involving descriptive statistics (e.g., mean, median, mode) or including the concept that statistics was used for gathering information about life. Accurate definition, but simplistic.	[When I hear 'statistics'] I think about math and lots of letters and numbers mixed together to create a formula. Hard. [What I expect to learn is] mean, median, mode, average, square roots. Statistics is a math class that collects numbers and uses them together.
3	Descriptive with variability	Participants' definitions included concepts of both descriptive statistics and the idea of variability. Accurate and complex definitions.	[What I expect to be covered in a statistics course:] how to collect data; how to interpret it using mathematics such as standard deviation; interpret and make graphs to understand the data. [Definition of "statistics":] a collection of data that can be manipulated [in] a mathematical manner to better understand its meaning. [A statistician would:] interpret and present data in a meaningful way.
4	Descriptive with Probability	Included clear concepts of probability. Some participants only mentioned probability while other participants explained concepts related to probability. Accurate, more complex definition.	(a) Statistics is the collection of data showing the probability of something happening. (b) [Statistics is] proportions, probabilities, and numbers that come together to form something you can put into words to better understand and perceive as something other than what it appears to be. (c) [Topics learned in the statistics course:] probability, proportions, categorical variables, quantitative variables, distribution.
5	Inferential	Participants' definitions included terminology or ideas related to inference. Concepts of *p*-value, random sampling, hypothesis testing, statistical significance, and confidence intervals were often mentioned in this category. Accurate definitions—these definitions were the most complex and comprehensive.	(a) [A statistician would] study the number of people/things which are likely similar and compar[e] them to others in different groups. (b) [When I think of statistics,] I think of means and proportions, and samples of populations all combined to form results that can be analyzed. (c) Statistics is the acquiring of information from a sample population to infer and analyze from.

Table 4. Conceptualization categories and quotes

Quantitative Value	Category Name[a]	Description	Illustrative Quote(s)
By focusing on techniques: (Gathering – Extrinsic Technical)			
1	Statistics is individual numerical activities	Participants discussed math equations and solutions such as graphing, problem solving, ratios, and fractions.	(a) math, solving problems, solving equations, word problems (b) Statistics is a way of figuring out mathematical equations and solutions.
2	Statistics is using individual statistics techniques	Participants mentioned statistics activities including limited mention of probability, sports statistics, and descriptive statistics.	(a) numbers, math, probability, an expensive calculator, and a huge load of homework (b) mean, median, mode, range
3	Statistics is a collection of statistical techniques	Responses in this category indicated that statistical techniques were connected to real life.	(a) [A statistician might] … find out how things are alike and differ from one another by studying them and putting them into groups. (b) [Topics covered in the statistics course:] Probability, test statistics, random sampling, point estimates (c) [Statistics is] the type of math which people will encounter in everyday life, (d) [A statistician] figures out probabilities and other math topics and analyzes the information related to them.
By focusing on data: (Applying – Extrinsic Meaning)			
4	Statistics is the analysis and interpretation of data	Definitions in this category went beyond just listing techniques and included the idea that statistics techniques lead to helpful interpretations	(a) [A statistician] would take data and make numbers into rational statements presenting information. (b) [A statistician could focus on] marketing and he/she finds out the number of what customers like or buy.
5	Statistics is a way of understanding real-life using different statistical models.	Responses in this category included the idea of interpretation and added statistical models. Participants often mentioned hypothesis testing, organization of data, and research design, with an emphasis on inference. Responses which mentioned z-scores were coded in this category.	(a) [I have learned] Probability Distribution, Hypothesis, z-score, Sample mean and sample proportion. (b) Statistics is the art and science of designing studies, analyzing results and data, translating data into knowledge and understanding involving design, describing and inferring.
By focusing on meaning: (Creating – Intrinsic Meaning)			
6	Statistics is an inclusive tool used to make sense of the world and develop personal meanings.	Participants conceptualized statistics as a tool that helped make sense of the world around them and even changed their understanding of life.	(a) A statistician probably creates surveys and compiles information at a professional level in order to inform many people about many different things. These statistics can influence important decisions and affect peoples' perspectives on different topics. (b) …statistics can help us understand more about our society (c) A statistician would find if there has been a growth in people who have a disease (or any other rise in numbers of people).

[a]The category names and hierarchy used are from Reid and Petocz (2002).

4.2. COMPARISON OF PRE-TEST AND POST-TEST DEFINITIONS AND CONCEPTUALIZATIONS

We next address the second research question, "Are there changes in student definitions and conceptualizations at the beginning as compared to their definitions toward the end of the course? If so, to what extent are these definitions different?" Because our content and concept scores are based on ordered categories we chose to use the sign test. The Content codes were given the numerical values listed in Table 3. Sixty-six percent (n = 31) were inaccurate or formative definitions of statistics. The median of the pre-course Content scores was 2 compared to a median of 5 for post-course Content scores, with post-course accuracy at 97% (n = 37). The sign test results indicated a significant improvement from pre- to post-score; in fact, only six participants' definitions remained at the same level while the remaining 31 participants improved their content understanding from pre to post (p-value < 0.01).

The Concept qualitative codes were converted to numerical scores by assigning the values listed in Table 4. The pre-course Concept median was 3 and the post-course Concept median was 4. Again, sign test results indicated significant improvement, with 31 students increasing their conceptualization understanding of statistics while a different six students remained the same (p-value < 0.01).

4.3. RELATIONSHIPS BETWEEN SATS, CONTENT, AND CONCEPTS

Mixed methods allowed the third and fourth research questions to be examined: "What is the relationship between student definition of statistics and attitudes toward statistics?" And "What is the relationship between student conceptualization of statistics and attitudes toward statistics?" Table 5 contains the Cronbach's alpha values for each of the six SATS-36 attitude components. The alpha values are in the acceptable range, with a minimum of 0.658 from the pre-SATS and a minimum of 0.724 from the post-SATS. The lower internal consistency on the pre-test for the Difficulty scale might be due to the lack of variability in the responses.

Table 5. Cronbach's alpha values for the SATS-36 attitude components

	Pre-SATS		Post-SATS	
Component (number of items)	Cronbach's α	*n*	Cronbach's α	*n*
Affect (6)	0.82	45	0.88	37
Cognitive Competence (6)	0.76	46	0.86	38
Value (9)	0.90	46	0.89	38
Difficulty (7)	0.66	47	0.72	37
Interest (4)	0.90	47	0.94	38
Effort (4)	0.93	46	0.88	38

Table 6. SATS Subscales

	Affect	Cognitive Competence	Value	Difficulty	Interest	Effort
Pre-Scores						
Mean	4.12	4.62	5.28	3.46	4.97	6.39
SD	1.10	0.96	1.02	0.70	1.19	1.14
n	45	46	46	47	47	46
Post-Scores						
Mean	4.00	4.63	4.82	3.38	3.69	5.98
SD	1.41	1.21	1.16	0.94	1.58	1.06
n	37	38	38	37	38	38
Mean Change	0.12	-0.01	0.46	0.08	1.28	0.41

Table 6 shows the pre- and post-subscores for the SATS-36 attitude components. The SATS scores represent averages from a 7-point Likert scale, with mean scores around 4 as neutral. With the exception of the Difficulty subscale, all five subscores have mean values above 4 on the pre-SATS. The Difficulty subscale can be thought of as perceived easiness, and thus, lower values should be interpreted as "statistics is a difficult subject" whereas higher values mean "statistics is easy." There was a drop in almost all of the subscales with Interest having the largest drop of 1.28 units.

SATS-36 and content of participants' definitions Assuming equal variances, the ANOVA found no significant difference in any of the mean SATS pre-scores among groups of participants in different Content categories (using an $\alpha = 0.002$ for an overall 0.05 level). Likewise, the post-SATS analysis did not reveal any significant differences in mean SATS subscales among groups of participants in different Content categories.

SATS-36 and conceptualizations Results of the ANOVA (assuming equal variances) suggest that there was no significant difference for the six subscales at an overall 0.05 level (using individual $\alpha = 0.002$) between the conceptualization levels, both in the pre- and post-tests.

Content of participants' definitions and conceptualizations Spearman's rho correlations resulted in significant correlation between pre-Content and pre-Concept of 0.451, $p < .01$. We found a weak, non-significant correlation between pre-Content and post-Concept of 0.350, $p = .034$; note that the Bonferroni correction required $p < .0125$ for significant results at the .05 level when four correlations were calculated. Students' pre-Concept and post-Concept ratings were also significantly correlated at 0.561, $p < .001$. However, participants' post-Content and post-Concept scores were not correlated. This showed that students' content knowledge of statistics was associated with how they conceptualized statistics at the beginning but not toward the end of the course.

5. DISCUSSION

The purpose of this project was to investigate students' perceptions of the field of statistics. Based on the results of this project, we now have a better understanding of students' content knowledge as well as their conceptualizations of statistics.

Tempelaar, Schim van der Loeff, and Gijselaers (2007) noted that students' attitudes and beliefs toward statistics were related to prior knowledge and reasoning abilities. We offer the possibility that student beliefs consist of their content knowledge and conceptual understanding of statistics and that these beliefs then impact student attitudes, although how these beliefs impact their attitudes remains a subject for future research. Collectively, beliefs and attitudes may comprise perceptions of statistics (See Figure 3).

Having presented a possibility for how these findings could work together to create an overall understanding of student perceptions of statistics, we now discuss the findings of this study in more detail. Even though 77% of the students in the sample had never had a statistics course, only 24% were at the lowest content understanding of statistics. However 55% were at the lowest conceptual understanding of statistics. The sign test results of Content and Concept scores confirmed that students' content understanding of statistics improved after taking a statistics course and also indicated that students' conceptualization of statistics became more complex. This would be expected due to learning. However, we also observed that the ceiling effect of the Concept scale was not as strong as the ceiling effect of the Content scale, as can be seen in Figures 1 and 2; that is, more students maximized their Content understanding than maximized their Concept understanding. Therefore, it appears that even students who can define statistics as having an inferential component may not conceptualize statistics as a way to make sense of the world. Even students who successfully complete a statistics course may not have developed Wallman's (2003) stated purpose of "statistical literacy."

Figure 1 shows that at the beginning of the course, most students responded in the third content category (Descriptive, 43%), which may indicate that although the majority of participants had an accurate understanding of statistics, most participants had only formative, rather than complex, understanding. Notice that toward the end of the course, most students thought of statistics as inferential statistics (59%). At the beginning of the semester, Figure 2 shows that most students thought statistics was about analysis and interpretation of data (26%), followed by using equations (23%), but by the time of the post-tests, the majority of the students realized the usefulness of statistics. Almost half (49%) of the participants perceived that statistics was about analysis and interpretation of data, and 27% thought that statistics was about understanding real-life using different statistical models. It seemed that students initially perceived statistics as useful, but equated statistics to mathematics.

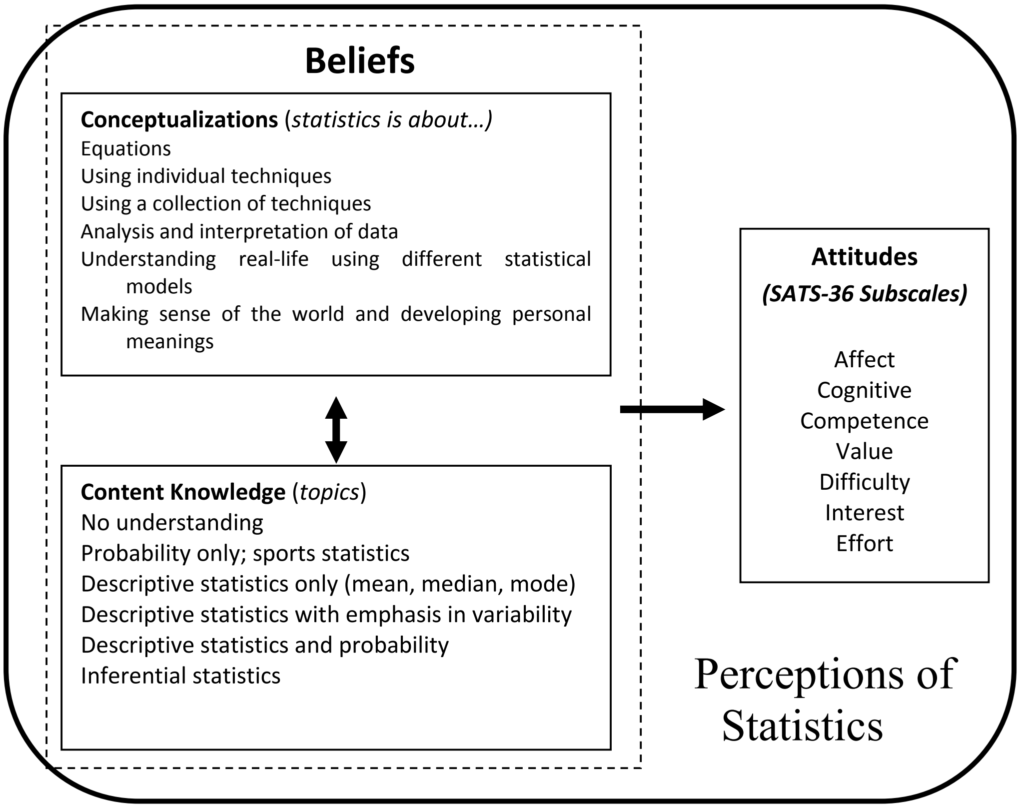

Figure 3. Perceptions of Statistics

Table 6 shows that the majority of SATS scores fit the pattern seen in previous research, that is, on average, attitudes stayed the same or had a small decrease for Affect, Cognitive Competence, and Difficulty while decreasing in magnitude for Value, Interest, and Effort (Bond, Schau, Pierce, & Schou, 2008; Schau & Emmioğlu, 2012). This decrease in mean attitude scores seems to indicate a less desirable attitude toward statistics at the end of a statistics course than at the beginning. It has been suggested that the reason students' attitudes toward statistics drops from pre-course to post-course is because students tend to choose the *neutral* answer of 4 on the SATS pre-survey if they did not know what "statistics" means in the wording of the item (Bond et al., 2008). Scores around 4, indicating neutral responses, have been common in the pre-course means of the SATS sub-scores. The SATS instructions guided participants to select "neither disagree nor agree" if one had no opinion.

Eleven students had a pre-Content score of 0 or 1 which was the lowest level of content understanding, and we might expect neutral responses in the range of 3.5 to 4.5 (not including the endpoints). To investigate this possibility, Figure 4 only reports the responses of the eleven participants with pre-Content scores of 0 or 1. Based on Figure 4, we observe that one student out of

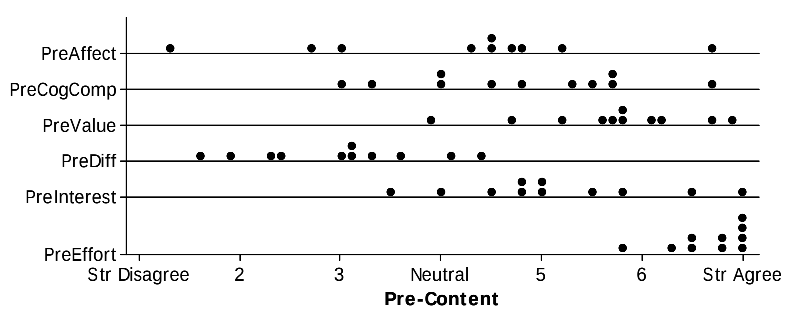

Figure 4. Eleven students had a level 0 or 1 on Pre-Content

eleven was in the neutral range for Affect and Value; two students out of eleven were in the neutral range for Cognitive Competence and Interest; three were in the neutral range for Difficulty; and zero were in the neutral range for Effort. Therefore, it appears that participants with a lower level of Content understanding did not tend to choose neutral responses on the pre-test.

At an overall level of $\alpha = 0.05$, we saw no significant difference in any of the SATS-36 subscale mean scores across the pre-Content categories and across the pre-Concept categories. This lack of significance is in keeping with the theory that students' Content and Conceptualization knowledge prior to taking the SATS-36 does not influence the SATS-36 subscales. However, during the exploratory data analysis phase of this study, we found that there might be differences among the pre-Content categories in the Difficulty subscale and among the pre-Concept categories in the Affect subscale from the pre-SATS. Figure 5 shows descriptively that the four students who had no understanding about statistics perceived statistics, on average, as more difficult than the other Content knowledge groups did. Figure 6 demonstrates descriptively that the three students who began the course at the highest conceptualization level had the highest levels of Affect. Note that these were descriptive rather than inferential conclusions, and that the inferential conclusion found no significant differences. More research is needed because these results are based on so few students.

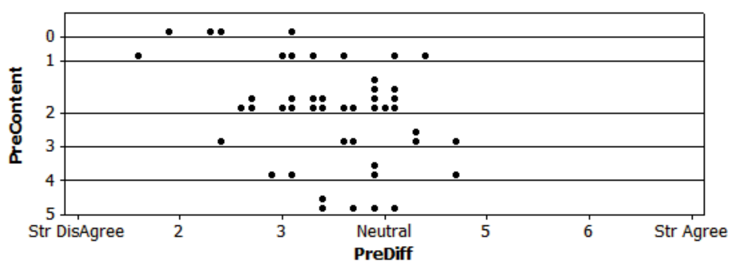

Figure 5. Pre-Difficult SATS Subscale at Pre-Content Levels

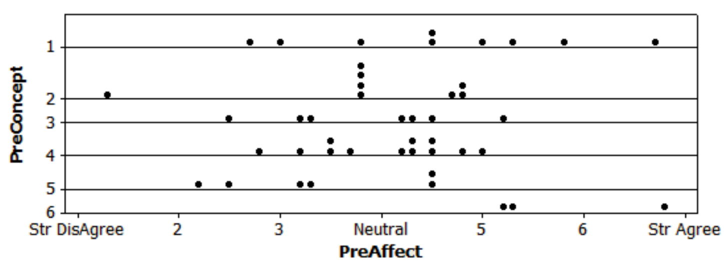

Figure 6. Pre-Affect SATS subscale at Pre-Concept Levels

In summary, we found that students had an overall accurate understanding of the word "statistics" even at the beginning of the course. In fact, the qualitative results of the pre-Content and pre-Concept surveys showed that the students had varying levels of understanding and conceptualization of statistics. The results from the post-Content and post-Concept surveys confirmed that students learned from the course. The pre- and post-analysis illustrated the growth of students' perception of statistics. Compared to previous studies, our results and the categories that we developed were similar to the study by Reid and Petocz (2002). Unlike Gordon (2004), we did not group the students according to their willingness to study statistics. However, we found that most of our students acknowledged the usefulness of statistics, which was similar to the majority of Gordon's students who were willing (not reluctant) to study statistics.

5.1. IMPLICATIONS FOR TEACHING AND RESEARCH

Teaching In contrast to the improvement in Content accuracy that was seen pre-course to post-course, participant improvement was seen at lower levels but not at higher levels in the Concept category. Perhaps this indicates that it was fairly easy for participants to grow in their understanding of the definition of statistics, but that participants' conceptualizations of the value or usefulness of statistics was more difficult to change. Even at the end of these statistics courses, participants' conceptualizations of statistics could improve. To help students develop statistical literacy (Wallman, 2003), instructors may need to concentrate on awareness of students' conceptualizations in addition to content knowledge. The most useful approach to teaching statistics may include helping students both understand and appreciate statistics by realizing that many students begin the course bringing in experience from mathematics courses. This study has contributed to our understanding that students' attitudes and beliefs (content knowledge and conceptualizations) about statistics are important variables that researchers, teachers, and statistics educators should consider.

Research In 1997, Gal, Ginsburg, and Schau inquired about how the "fuzziness" of the word "statistics" could affect the validity of attitude surveys. Our results demonstrated that students were familiar enough with the word to participate in attitude surveys (at least with the SATS). Although we found no significant difference in any of the mean SATS subscores across either Content or Concept groups, we think that further research could examine possible connections between the Difficulty subscale and low pre-Content scores. Perhaps students who have more complex definitions of statistics also perceive statistics as being more (or less) difficult than do other students. A better understanding of this connection between definition and perception of difficulty could help explain whether students' attitudes toward statistics become less favorable as they better understand the complexity of what statistics entails. Future research could also explore a connection between the Affect subscale and high pre-Concept scores. Perhaps students with high Concept scores, which would indicate that they recognize that statistics can be used to make sense of the world, also have more positive attitudes toward statistics.

The trend of SATS scores declining from pre- to post-course assessment was repeated in this study. However, the authors wondered whether asking students about their attitudes toward statistics immediately upon completion of a difficult course may negatively impact attitudes. At the end of the semester, students' attitudes may be influenced by worrying about grades and being overwhelmed by preparing for final exams. It is possible that assessing students' attitudes toward statistics a few months after a statistics course was completed could allow for a more accurate assessment of attitude changes. Future research could examine this possibility by doing follow-up studies using the SATS several months after a statistics course. Additionally, one wonders whether this drop in attitudes is unique to statistics. Could it be that this phenomenon occurs in other disciplines as well?

The *Perceptions of Statistics* short-answer survey should undergo additional evaluation; however, this tool supplied a rich source of data from which the authors were able to obtain content and conceptualization classification without the expense and time of personal interviews. If the coding schemes developed in this article and by Reid and Petocz (2002) are used in future research, additional studies could explore the possibility of creating formalized definitions of each category. Researchers could decide whether gathering students' own perceptions would be a valuable use of resources or whether a quantitative survey or assessment could be used to assess participants' understanding more efficiently.

In starting to better understand beliefs toward statistics, we agree with Tempelaar, Schimvan der Loeff, and Gijselaers (2007) in encouraging researchers to continue focusing on the construct of student beliefs about statistics. It is possible that beliefs entail more components than are included in the framework shown in Figure 3. For example, could common knowledge and attitudes toward mathematics influence attitudes and perception of statistics? It is also possible that a better understanding of how these components impact student perceptions of statistics may provide more options for instructors to intervene with students in ways that will increase their chances of using statistics in their futures.

5.2. STRENGTHS AND LIMITATIONS OF THIS STUDY

Strengths of this study included the focus on student perceptions which allowed participants' own definitions and conceptualizations to be better understood. Gathering very focused information from participants reduced the demand on participant time from what would have been required for interviews and also provided adequate qualitative information to answer our research questions. This study wove together themes from several definitions of statistics in a way that we believe will be a valuable contribution for future research. Additionally, participants' responses allowed us to build on previous research findings by using Reid and Petocz's (2002) categories for conceptualization of statistics. Finally, examining the connection between participants' qualitative information and SATS quantitative information for attitudes was also a strength of this study.

Weaknesses included the small number of participants in this study and the limited number of courses from which participants were recruited. Additionally, although we were excited to consider such a wide variety of variables in the content, conceptualization, and SATS subscales, the small sample size may have resulted in Type II errors.

6. CONCLUSION

This study was designed to help statistics educators and researchers better understand students' definitions of statistics. The qualitative findings provided descriptions of the way students conceptualize statistics and the content of what they think statistics is. The mixed methods design of this study allowed quantitative analysis to include both the qualitative findings and SATS-36 scores. Results led to a broad understanding of student perceptions of statistics, with perceptions including beliefs and attitudes. These findings can be used to guide instruction of statistics and future research.

ACKNOWLEDGEMENTS

The research reported in this paper was supported partially by the Consortium for the Advancement of Undergraduate Statistics Education (CAUSE), through NSF Award #0618790. We are grateful to for the organizers of the USCOTS 2009 (United States Conference On Teaching Statistics) for starting our Attitudes Research Cluster. Although none of her research appears or was used in this paper, the authors acknowledge the undergraduate student, Gloria Lehr, who helped gather pilot survey data. The authors are grateful for the comments from the editors and referees. These comments greatly improved the article.

REFERENCES

Ben-Zvi, D., & Garfield, J. B. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3-15). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Bond, M. E., Schau, C., Pierce, R. L., and Schou, S. (2008). Research using the *Survey of Attitudes Toward Statistics* (SATS). *Proceedings of the Section on Statistical Education, 2008 Joint Statistical Meetings* (pp. 3750–3763). Alexandria, VA: American Statistical Association.

Carlson, K. A., & Winquist, J. R. (2011). Evaluating an active learning approach to teaching introductory statistics: A classroom workbook approach. *Journal of Statistics Education, 19*(1), 1–23.
[Online: http://www.amstat.org/publications/jse/v19n1/carlson.pdf]

Carnell, L. J. (2008). The effect of a student-designed data collection project on attitudes toward statistics. *Journal of Statistics Education, 16*(1).
[Online: http://www.amstat.org/publications/jse/v16n1/carnell.html]

Chiesi, F., & Primi, C. (2010). Cognitive and non-cognitive factors related to students' statistics achievement. *Statistics Education Research Journal, 9*(1), 6–26.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ9%281%29_Chiesi_Primi.pdf]

Cobb, G.. (1992). Teaching statistics. In L. A. Steen (Ed.) *Heeding the call for change: Suggestions for currricular action* (MAA Notes No. 22, pp. 3–43). Washington, DC: Mathematical Association of America.

Cobb, G., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly, 104*(9), 801–823.

Coetzee, S., & Van der Merwe, P. (2010). Industrial psychology students' attitudes towards statistics. *SA Journal of Industrial Psychology, 36*(1). [Online: http://www.sajip.co.za/index.php/sajip/article/viewPDFInterstitial/868/873]

Creswell, J. W. (1998). Qualitative inquiry and research design: Choosing among five traditions. Thousand Oaks, CA: Sage.

Dauphinee, T. L., Schau, C., & Stevens, J. J. (1997). *Survey of Attitudes Toward Statistics*: Factor structure and factorial invariance for women and men. *Structural Equation Modeling, 4*(2), 129–141.

delMas, R. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 79–95). Dordrecht, The Netherlands: Kluwer Academic Publishers.

delMas, R. (2011, May). *It takes a village: Future directions for statistics education research*. Paper presented at the USCOTS 2011: The Next BIG Thing, Raleigh-Durham, NC.

Dempster, M., & McCorry, N. K. (2009). The role of previous experience and attitudes toward statistics in statistics assessment outcomes among undergraduate psychology students. *Journal of Statistics Education*, *17*(2). [Online: http://www.amstat.org/publications/jse/v17n2/dempster.html]

DeVeaux, R. D., & Velleman, P. F. (2008). Math is music; statistics is literature (Or, why are there no six-year-old novelists). *AMSTATNEWS*, 54–58.

Evans, B. (2007). Student attitudes, conceptions, and achievement in introductory undergraduate college statistics. *The Mathematics Educator*, *17*(2), 24–30. [Online: http://math.coe.uga.edu/tme/issues/v17n2/v17n2.pdf]

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 37–51). Amersterdam: IOS Press. [Online: http://www.stat.auckland.ac.nz/~iase/publications/assessbk/chapter04.pdf]

Gal, I. (2002). Adults' statistical literacy: meanings, components, responsibilities. *International Statistical Review, 70*(1), 1–51.

Garfield, J. B., & Gal, I. (1999). Teaching and assessing statistical reasoning. In L. V. Stiff & F. R. Curcio (Eds.), *Developing mathematical reasoning in grades K-12: 1999 Yearbook.* (pp. 207–220). Reston, VA: National Council of Teachers of Mathematics.

Gordon, S. (2004). Understanding students' experiences of statistics in a service course. *Statistics Education Research Journal, 3*(1), 40–59. [Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ3%281%29_gordon.pdf]

Groth, R. E. (2010). Situating qualitative modes of inquiry within the discipline of statistics education research. *Statistics Education Research Journal, 9*(2), 7–21. [Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ9%282%29_Groth.pdf]

Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco, CA: Jossey-Bass.

Higgins, J. J. (1999). Nonmathematical statistics: A new direction for the undergraduate discipline. *The American Statistician, 53*(1), 1–6.

Hilton, S. C., Schau, C., & Olsen, J. A. (2004). *Survey of Attitudes Toward Statistics*: Factor structure invariance by gender and by administration time. *Structural Equation Modeling, 11*(1), 92–109.

Kalinowski, P., Lai, J., Fidler, F., & Cumming, G. (2010). Qualitative research: An essential part of statistical cognition research. *Statistics Education Research Journal, 9*(2), 22–34. [Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ9%282%29_Kalinowski.pdf]

Kazdin, A. E. (1998). *Research design in clinical psychology* (3rd ed.). Boston: Allyn and Bacon.

Maxwell, J. A. (1998). Designing a qualitative study. In L. Bickman & D. J. Rog (Eds.), *Handbook of applied social research methods* (pp. 69–100). Thousand Oaks, CA: Sage.

Merriam, S. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass Publishers.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Newbury Park, CA: Sage.

Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review, 65*(2), 123–165.

Patton, M. Q. (2005). Qualitative research. In B. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioral science*. New York: Wiley & Sons.

Posner, M. A. (2011). The impact of a proficiency-based assessment and reassessment of learning outcomes system on student achievement and attitudes. *Statistics Education Research Journal, 10*(1), 3–14.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ10%281%29_Posner.pdf]

Reading, C., & Shaughnessy, J. M. (2004). Reasoning about variation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 201–226). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Reid, A., & Petocz, P. (2002). Students' conceptions of statistics: A phenomenographic study. *Journal of Statistics Education, 10*(2).
[Online: http://www.amstat.org/publications/jse/v10n2/reid.html]

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. Elliott (Eds.), *Thinking and reasoning with data and chance* (pp. 323–334). Reston, VA: National Council of Teachers of Mathematics.

Rossman, G. B., & Rallis, S. F. (2003). *Learning in the field: An introduction to qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

Schau, C. (1992). *Survey of Attitudes Toward Statistics* (SATS-28).
[Online: http://evaluationandstatistics.com/]

Schau, C. (2003a). *Survey of Attitudes Toward Statistics* (SATS-36).
[Online: http://evaluationandstatistics.com/]

Schau, C. (2003b, August). *Students' attitudes: The "other" important outcome in statistics education*. Paper presented at the Joint Statistical Meetings, San Francisco, CA.
[Online: http://evaluationandstatistics.com/JSM2003.pdf]

Schau, C., & Emmioğlu, E. (2012). Do introductory statistics courses in the united states improve students' attitudes?. *Statistics Education Research Journal*, *11*(2), 70–79.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ11%282%29_Schau.pdf]

Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the *Survey of Attitudes Toward Statistics*. *Educational and Psychological Measurement, 55*, 868–875.

Sorge, C., & Schau, C. (2002, April). Impact of engineering students' attitudes on achievement in statistics. Paper presented at the *American Educational Research Association*. New Orleans, LA.
[Online: http://evaluationandstatistics.com/AERA2002.pdf]

Tempelaar, D. T., Gijselaers, W. J., Schim van der Loeff, S., & Nijhuis, J. F. H. (2007). A structural equation model analyzing the relationship of student achievement motivations and personality factors in a range of academic subject-matter areas. *Contemporary Educational Psychology, 32*(1), 105–131.

Tempelaar, D. T., Schim van der Loeff, S., & Gijselaers, W. H. (2007). A structural equation model analyzing the relationship of students' attitudes toward statistics, prior reasoning abilities, and course performance. *Statistics Education Research Journal, 6*(2), 78–102.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ6%282%29_Tempelaar.pdf]

Utts, J. (2003). What educated citizens should know about statistics and probability. *The American Statistician, 57*(2), 74–79.

Vanhoof, S., Kuppens, S., Sotos, A. E. C., Verschaffel, L., & Onghena, P. (2011). Measuring statistics attitudes: Structure of the *Survey of Attitudes Toward Statistics* (SATS-36). *Statistics Education Research Journal, 10*(1), 35–51.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ10%281%29_Vanhoof.pdf]

Wallman, K. S. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association, 88*(421), 1–8.

Watson, J. M. (1997). Assessing statistical thinking using the media. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education*. Amsterdam: IOS Press.
[Online: http://www.stat.auckland.ac.nz/~iase/publications/assessbk/chapter09.pdf]

Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal, 2*(2), 3–46.
[Online: http://www.stat.auckland.ac.nz/~iase/serj/SERJ2%282%29_Watson_Callingham.pdf]

MARJORIE E. BOND
Department of Mathematics and Computer Science
Monmouth College
Monmouth, IL 61462