PAINT-BY-NUMBER OR PICASSO? A GROUNDED THEORY PHENOMENOGRAPHICAL STUDY OF STUDENTS' CONCEPTIONS OF STATISTICS

NICOLA JUSTICE
Pacific Lutheran University
njustice@plu.edu

SAMANTHA MORRIS
Pacific Lutheran University
morrissa@plu.edu

VERONIQUE HENRY
University of Colorado, Colorado Springs
vhenry@uccs.edu

ELIZABETH BRONDOS FRY
Saint Catherine University
ebfry704@stkate.edu

ABSTRACT

Statistics students' conceptions of the work of statisticians and the discipline of statistics may play an important role in the topics to which they attend and their interest in pursuing further study. To learn about students' conceptions, we collected open-ended survey responses from 44 undergraduate students who had completed introductory statistics courses. We used a grounded theory phenomenographical qualitative approach to identify several themes in students' conceptions. In addition to the test-and-procedure conception, we offer several other themes, such as acknowledgement of variation and the role of ethical integrity. We use a metaphor of painting styles to compare to experts' conceptions of statistics. By identifying "seeds" of what may be developed into expert conceptions, these preliminary results set possible foundations to explore trajectories that may help shape students' conceptions of statistics.

Keywords: *Statistics education research; Students' experiences; Phenomenography; Variation*

1. INTRODUCTION

There is no question as to the data-driven nature of our society and consequently employers' demand for professionals who can manage and analyze data. Projections predict shortages of thousands of employees with skills to make sense of data (e.g., Manyika et al., 2011). Wild et al. (2018) posit that there is a need to (1) equip statistics students at all levels to reach various competencies managing and interpreting data, and (2) attract and recruit more students into the discipline who can become responsible, innovative producers of data-based statistical results.

Regarding the first need, equipping students to reach various statistical competencies is certainly not easy. At introductory levels, delMas et al. (2007) report that undergraduate students who have completed an introductory statistics course struggle with many basic statistical concepts (e.g., correctly interpreting *p*-values). In a more recent study, Fry (2017) suggests that undergraduate students in introductory statistics courses struggle to distinguish between the implications of random sampling versus random assignment, even after participating in focused, research-based teaching modules on the implications of study design for appropriate conclusions. Researchers have explored potential reasons for why statistics is so difficult to understand. Tversky and Kahnneman (1974) identify paradoxes in

Statistics Education Research Journal, 19(2), 76–102, http://iase-web.org/Publications.php?p=SERJ
© International Association for Statistical Education (IASE/ISI), June, 2020

the foundations of probabilistic thinking that make it difficult for humans to make sense of chance and uncertainty. Lexical ambiguities may also contribute to students' confusion (e.g., Dunn et al., 2016; Kaplan et al., 2009).

Teachers may be able to more easily facilitate growth in students' understanding of statistics if they can anticipate the conceptions of statistics that students bring with them into the classroom. Research in cognitive psychology suggests that it is essential for teachers to be aware of and build on such conceptions. Bransford et al. (2004) explain how learning is enhanced when teachers "pay attention to the knowledge and beliefs that learners bring to a learning task, use the knowledge as a starting point for new instruction, and monitor students' changing conceptions as instruction proceeds" (p. 12).

A greater understanding of students' conceptions of statistics may also help with the second need identified by Wild et al. (2018), namely, to attract and recruit more students into the discipline. If students have incomplete or misleading understandings of the discipline, many ideal candidates may be less interested in taking more courses or pursing statistics degrees. Should common false conceptions of statistics be identified and targeted for change, potential recruits may feel more interested in exploring the discipline further.

There are many examples of efforts to try to combat misconceptions about statistics. The American Statistical Association (ASA) launched a public education campaign, *This is Statistics* (http://thisisstatistics.org), designed to help learners of all ages come to a better understanding of the discipline. Statistics enthusiasts and experts such as Holcomb (2015) participate in outreach efforts, offering to students ideas of why they enjoy being statisticians. These efforts to change students' conceptions of statistics are evidence that students' conceptions of statistics are worthy of study.

This preliminary study explores conceptions of statistics held by undergraduate students who had completed an introductory statistics course. A phenomenographic (e.g., Marton, 1981) grounded theory (e.g., Creswell, 2007) qualitative approach was used to collect data and identify new possible aspects of students' understanding of the discipline of statistics. Results are used to compare students' views to experts' views of the discipline and offer an example framework for students' conceptions. We use the metaphor of different artistic approaches (e.g., paint-by-number, Picasso) to illustrate the different conceptions.

2. BACKGROUND

As background for exploring students' conceptions of statistics, we begin with a review of experts' conceptions of statistics and how statistics may be compared to mathematics. The focus is based primarily on seminal work by Pfannkuch and Wild (2000). Then, we review previous research on students' conceptions of statistics and establish this study's unique contribution.

Although the focus of this paper is on conceptions, here we do not attempt the daunting task of defining the term *conceptions*, or distinguishing it from its many aliases (e.g., *beliefs*, *attitudes*, *values*, *opinions*, *perceptions*, *preconceptions*, *dispositions*, *perspectives*, and *knowledge*). For attempts to define and distinguish these concepts, we refer readers to experts such as Pajares (1992).

Marton (1981), a founder of the phenomenographical approach used in this study, uses *conceptions* along with the terms *descriptions*, *ways of understanding*, and "people's ideas about the world (or their experience of it)" (p. 178). Following the model of Marton, we will henceforth use the term *conceptions* to represent individuals' interpretive descriptions of the world. Our use of the term *conceptions* goes beyond *perceptions* and *experiences* into the realm of *expressed* descriptions of those experiences. We do not distinguish between conceptions that are new or naïve versus those that are established and well thought out (Marton, 1994). We do, however, break from convention by evaluating the extent to which student conceptions appear to match those of experts. We focus on what Rokeach (1968) may describe as statements that are descriptive (e.g., statistics involves calculations), rather than prescriptive (e.g., I must learn more statistics) or evaluative (e.g., statistics is fun).

2.1. EXPERTS' CONCEPTIONS OF STATISTICS

In the first chapter of the *International Handbook of Research in Statistics Education*, Wild et al. (2018) scratch the surface of the difficult and often controversial question, "What is Statistics?" They draw from several sources, including a definition offered by Davidian and Louis (2012), which has since been adopted by the American Statistical Association for various publications and resources: "Statistics is the science of learning from data, and of measuring, controlling for, and communicating uncertainty" (p. 12). Wild et al. add that statistics is a "meta-discipline that *thinks about how to think* about turning data into real world insights" (2018, p. 7). It is clear that statistics has no consensus definition. It generally can be agreed, however, that the discipline involves using data to grapple with uncertainties involved with real-world problems and decision-making.

Pfannkuch and Wild (2000) describe interviews with six professional applied statisticians designed to examine statistical thinking. They offer a four-dimensional framework (Figure 1), which is useful for identifying the skills, knowledge, processes, and dispositions necessary for conducting professional statistical work. Dimension 1 involves participation in the investigative Problem-Plan-Data-Analysis-Conclude (PPDAC) cycle. The second dimension includes what the authors call *types of thinking*, which are divided into general types and types fundamental to statistical thinking. The latter group includes recognition of the need for data, transnumeration, consideration of variation, reasoning with statistical models, and integrating the statistical and contextual. The third dimension involves participation in an interrogative cycle, which includes generating, seeking, interpreting, criticizing, and judging ideas. Finally, the fourth dimension includes dispositions a statistician must have, including curiosity, logical thinking, and perseverance. Wild and Pfannkuch note that variation is central to statistical thinking.

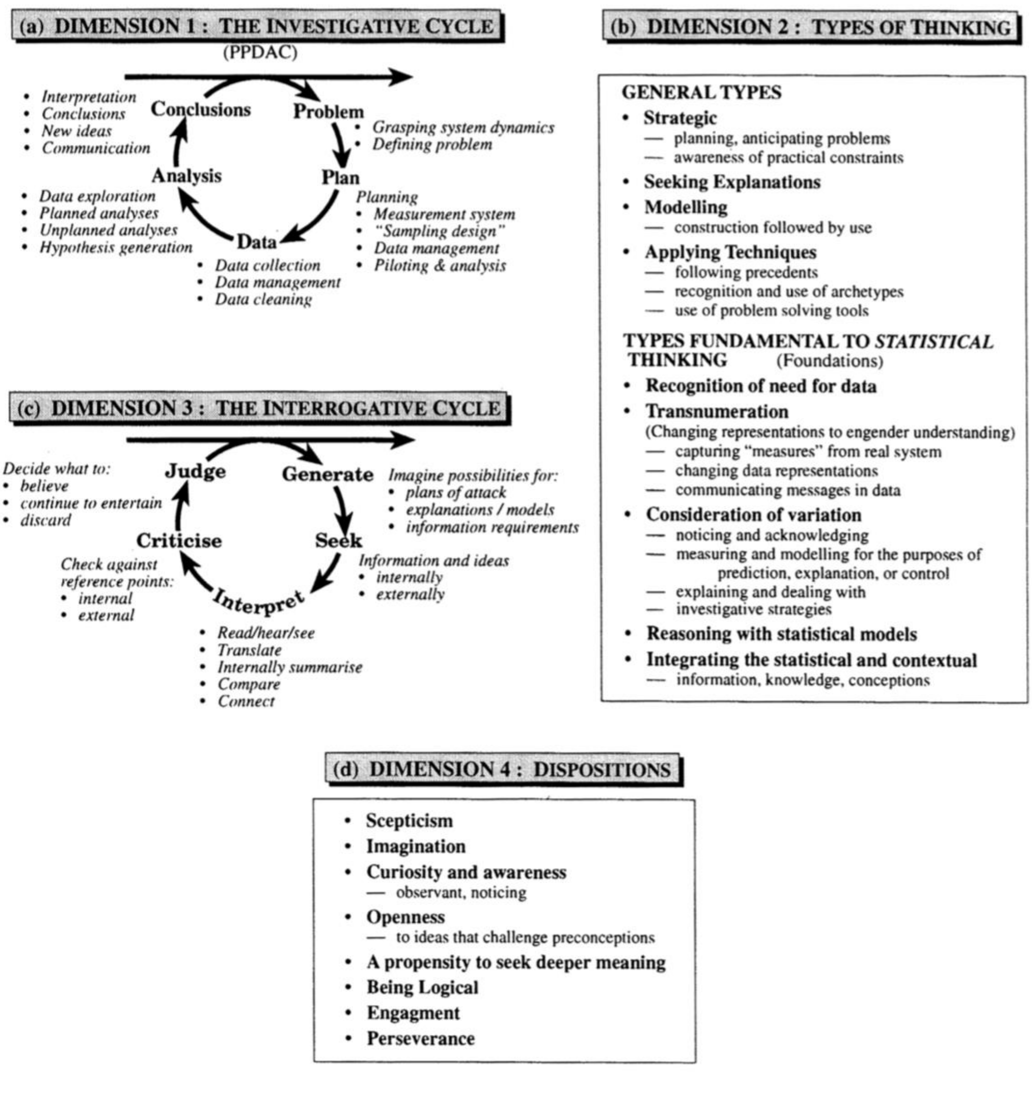

Figure 1. Framework for statistical thinking. Reprinted from "Statistical thinking in empirical enquiry," by Wild & Pfannkuch (1999), *International Statistical Review, 67*(3), p. 226. Reprinted with permission.

As noted by Wild et al. (2018), the expert conceptions of statistical thinking are more complex than what can reasonably be expected from novice statistical thinkers. Experts in early statistics education (elementary and secondary level) recommend exposing students to statistical thinking at gradually more sophisticated levels (e.g., Franklin et al., 2007). One line of research more akin to novice understandings of statistics focuses on *informal statistical inference* (ISI). In contrast to formal statistical inference, ISI has three qualities: (1) claims or generalizations that go beyond the sample data, (2) evidence (in particular, data-based evidence) to support the claim, and (3) probabilistic language to express the uncertainty of the claims and generalizations (Makar & Rubin, 2009). Zieffler et al. (2008) offer a framework for ISI, suggesting it involves (1) judgements about populations based on a sample, (2) drawing upon prior knowledge, and (3) articulating evidence-based arguments for the judgements. A commonality in these definitions is the act of making evidence-based claims in light of uncertainty.

Statistical experts appear to agree that statistical thinking is distinctive from mathematical thinking. Cobb and Moore (1997) point out how context is often treated as irrelevant in mathematics whereas it is critical to statistical analysis and conclusions. The authors also suggest that statistics differs from mathematics in problem-solving approach: Whereas "mathematical theorems are true; statistical methods are sometimes effective when used with skill" (Cobb & Moore, 1997, p. 810). Rossman et al. (2006) agree that statistics is much more context-driven and results in conclusions that are less certain than in mathematics.

Although experts agree these are important distinctions between statistics and mathematics, there is evidence that success in the subjects may be related. After evaluating the role of mathematical skills for 124 undergraduate statistics students, Primi et al. (2016) suggest that mathematics proficiency is related to student success in introductory statistics courses.

2.2. STUDENTS' CONCEPTIONS OF STATISTICS AND LEARNING STATISTICS

To date, the studies conducted to explore students' conceptions of statistics tend to take a phenomenographical qualitative approach (e.g., Reid & Petocz, 2002; Gordon, 2004), a method of qualitative analysis that seeks to identify a variety of human experiences of a particular phenomenon. Perhaps the most cited study is by Reid and Petocz (2002), who use the phenomenographical approach to collect and analyze interview data from 20 statistics students. Their resultant six-part, hierarchical framework has at the lowest end views of statistics as isolated or integrated tests and procedures; at the highest levels a way of understanding and making sense of the real world. Bond et al. (2012) come to similar conclusions after analyzing open-ended survey data using a "loose approach," in which they sought to have the results emerge from the data rather than a pre-conceived framework. Bond and colleagues conclude that their emergent categories were akin to the hierarchy offered by Reid and Petocz. Gordon (2004) also uses a phenomenographical approach to explore psychology students' conceptions of statistics. She does not report a hierarchical set of categories, however, most of her categories are related to those offered by Reid and Petocz. For example, the three categories (1) *No Meaning*, (2) *Processes* or algorithms, and (3) *Mastery* of statistical concepts and methods can be loosely tied to Reid and Petocz's tests and procedures end of the spectrum. Gordon's final two categories, (4) a *Tool* for getting results in real life and (5) *Critical Thinking*, can be loosely tied to Reid and Petocz's categories of tool for understanding the real world and making meaning, respectively. Özmen and Baki (2018) arrive at similar categories from Turkish students and report the addition of two more categories: *research and surveys* and *terminology and probability concepts*. The former is difficult to distinguish from Reid and Petocz's ideas of tools used for answering real life questions. In summary, the results of studies of students' conceptions of statistics appear to be very similar to one another, with a general theme of students' conceptions falling along one (hierarchical) spectrum: tests, procedures, and manipulation of numbers versus meaningful, useful, and impactful real-world applications.

In another phenomenographical study of interviews with n = 12 students who self-identified as having negative attitudes about statistics, Hedges and Harkness (2017) explore a slightly different aspect of students' conceptions of statistics: the relationship between statistics and mathematics. They offer three categories of students' conceptions: (1) statistics is mathematics, (2) statistics is like geometry (both are types of math; both are "nit-picky"), and (3) statistics is almost not mathematics (e.g., disappointment and fear that it is *more* than working with numbers). The authors note that students who felt there is a difference between the subjects still struggle to articulate the nature of the differences.

Instead of studying students, Findley and Kaplan (2018) explore conceptions held by novice teachers (graduate student instructors) by collecting data from interviews, surveys, classroom observations, and other teaching artifacts. Preliminary findings suggest that the novice teachers' conceptions about statistics can be classified according to three spectra: (1) statistics is driven by theory vs. application, (2) doing statistics is methodical vs. flexible, and (3) whether knowledge of mathematics is critical to learning statistics.

2.3. SUMMARY

Much has been written about the nature of statistics, experts' conceptions of statistical thinking, and how statistics is distinctive from mathematics. Experts' conceptions tend to include several dimensions. A common theme among them is the prominent role of variation and uncertainty. Studies of students' conceptions typically focus on one dimension, namely tests and procedures versus real-world applications. Although students may not hold expert-level depth of understanding, it is reasonable that students, like experts, may hold multidimensional views of the discipline. Current research is beginning to identify other possible candidates for aspects of graduate students' conceptions of statistics, but as yet does not explore additional dimensions of conceptions held by students at the introductory level.

In addition, there is a need to explore connections between students' and experts' conceptions of statistics. If teachers are charged with the task of building on students' previous conceptions, it would be helpful to supply teachers with suggestions of what their students' previous conceptions may be. Moreover, research that offers links between students' and experts' conceptions may help teachers and future researchers identify learning trajectories for shaping students' conceptions of statistics to appropriately reflect those of experts.

This study seeks to explore possible additional aspects of students' conceptions of statistics. Specifically, the research questions are

(1) What are aspects of students' conceptions of statistics and statistical thinking?

(2) How do students' conceptions of statistics and statistical thinking reflect experts' conceptions?

We approach this research from the perspective of statistics education researchers, and our responses to the research questions are framed primarily by Wild and Pfannkuch's (1999) account of experts' conceptions of statistical thinking.

3. METHODS

The qualitative approach of this study was chosen by the process of elimination after reviewing and considering several candidate approaches (e.g., Creswell, 2007; Saldaña, 2015). As is often the case, it was appropriate to blend grounded theory with another approach, namely phenomenography. These two approaches have many similarities, but a few nuanced differences rendered it appropriate to draw from both. We discuss the rationale for our blended approach in Section 3.1.

Principles of phenomenographical and grounded theory approaches drove much of our methodology. With the phenomenographic goal of identifying variation in conceptions, we gathered data with open-ended survey questions from 44 students who had experienced a large variety of introductory statistics course instructors. As is consistent with both phenomenography and grounded theory, data were analyzed using an iterative, collaborative process, to arrive at results that include both a set of themes and a framework identifying the most structurally substantive differences that emerged from the data (Marton, 1986).

3.1. THEORETICAL FRAMEWORK

The motivation for the study, data collection, and analysis were all guided by the belief that students' ideas of reality are constructed based on previous knowledge through assimilation and accommodation of ideas into new and existing schemas (e.g., Piaget, 1967) and so their conceptions of statistics are foundational to what knowledge they will acquire. Moreover, instructors must encourage students to build on prior knowledge. To this end, one primary purpose of this study was to learn about students' conceptions of statistics.

With this primary purpose in mind, our theoretical framework naturally draws from the second-order perspective of phenomenography. Whereas a first-order perspective seeks to describe various aspects of the world, a second-order perspective (and the perspective that is taken here) seeks to describe peoples' *experiences* of various aspects of the world. This perspective considers conceptions worthy of study—independent of the truth. Marton (1986) dubs studies of such nature as *phenomenography*, which involves the study of the "apprehended (experienced, conceptualized)"…. "content-oriented and interpretive descriptions of the qualitatively different ways in which people perceive and understand their reality" (Marton, 1981, p. 1).

As is common with phenomenography, the focus of our research is on the *variation* in the ways people experience the phenomenon of interest. The focus of the variation, however, is not within individuals but rather on the "collective intellect" of many (Marton, 1981, p. 1). The collection of categories of conceptions is considered stable, even if individuals move across categories over time.

In addition to a phenomenographical approach, our study lends itself to grounded theory, which goes beyond description to seek to show action and change (Strauss & Corbin, 1990). This is fitting because our motivations are to build foundations for a body of research on actionable approaches for shaping students' conceptions. Moreover, our goals are to conduct preliminary work that can be used to compare students' conceptions to experts' conceptions of statistics, perhaps for the purposes of designing interventions to help shape students' conceptions. This goal is also aligned with grounded theory, which often seeks to discover foundations for future areas of research. In light of these goals, we blended our phenomenographical approach with grounded theory.

We depart from the typical phenomenography assumptions that the different conceptions are logically related to one another by way of *hierarchical* inclusive relationships (Marton & Booth, 1997). This is in part due to the fact that many phenomenographies study conceptions of natural phenomena (for example forces of nature), whereas our study is of conceptions of a discipline, which is socially constructed. Therefore, we do not expect there to be the same tight hierarchical structure of conceptions. We do look for logical structure relating the different meanings, but also allow for our results to take form in a non-hierarchical outcome space.

We will also depart from phenomenography by reporting *themes*—as opposed to categories—that emerge from the data. Our rationale is that in this preliminary study we wish to avoid inappropriate nuances of mutual exclusiveness, exhaustiveness, and discrete cut-points that are associated with the term *category* and that cannot be supported by our data. Therefore, rather than categories (which are common in phenomenography), we opt instead to present our results as themes, and reserve for future research exploration of potential cut points.

It follows that the results we seek will be reported as an outcome space that only partially aligns with that of phenomenography. Marton and Booth's (1997) judgement of the quality of the phenomenographical outcome space includes the following: (1) each category represents something distinctive, (2) categories are logically related (typically hierarchical), and (3) outcomes are parsimonious. We hold to the first and the third of these criteria, seeking a parsimonious set of distinctive categories that describe conceptions of the discipline of statistics. We allow for our results to take a non-hierarchical outcome space, however, as is appropriate for grounded theory, and which may help our goals of comparison to experts and potential action.

3.2. PARTICIPANTS

The participants were students at a small, private liberal arts university located in the Northwestern United States. Human subjects approval was granted by the host institution, and pseudonyms are used in this report. All participants had completed—or nearly completed—an introductory statistics course.

Two groups of participants were solicited. Group One (n_1 = 26) was composed of students in an upper-level Probability and Statistical Theory course. The course is an elective for students majoring in mathematics and computer science, and most of the participants in Group One were represented by those majors. The course attracted a few other majors as well (e.g., physics; psychology), as it is one of the few possible electives to be chosen by students earning a minor in statistics.

Data from students in Group One were collected on the first day of class, before the course began. Although the course had an introductory statistics course prerequisite requirement, recent personnel changes at the institution meant that the prerequisite requirement was met for different students by many different instructors. As is aligned with phenomenographic research, herein was a particular opportunity to solicit more variation than might typically be expected at a small liberal arts college with limited statistics course offerings; there was not just one pipeline nor one instructor by which the students came. About 30% of participants in Group One took an introduction to mathematical statistics course taught by the first author using the *ISCAM* curriculum (Chance & Rossman, 2017), which uses simulation-based and traditional methods to teach statistical inference via investigations of real-world problems.

Participants in Group Two ($n_2 = 18$) were near completion of an introductory biostatistics course. Nearly all of the students in this group were biology (including pre-med) majors, although a few other majors (e.g., chemistry, latino studies) were represented. Participants in Group 2 were taught using Lock et al. (2013), which uses simulation-based and traditional methods to teach statistical inference.

To form the final data set, data from Group One and Group Two were merged. As all participants had completed or nearly completed introductory statistics, for this study no distinction was made between these groups.

3.3. SURVEY INSTRUMENT

Data were collected via an online survey. As this study is motivated largely by the need for more statistically-equipped graduates entering the workforce, the first half of the survey focused on what a statistician does and the traits and skills a professional statistician needs. The second half of the survey was designed to examine students' conceptions of statistics itself. The survey was designed to ask broad, open-ended questions so as to elicit more diversity of conceptions of statistics.

The instrument was created via an iterative process including consultation with several experts in statistics education assessment. First, the first author wrote several questions intended to elicit students' views about statistics and its relationship to mathematics. Because a primary purpose of this study was to develop students' conceptions grounded in the data, the items were designed to be open-ended and not given to leading students in any particular direction. The exceptions were the fifth and sixth questions, which asked students to identify the differences between statistics and mathematics and statistics and other natural sciences, respectively. These items were phrased to imply a distinction so as to solicit how students differentiate statistics from other science fields.

A draft of the survey was reviewed by an independent statistics education researcher from another institution in a different region of the United States. This colleague was solicited because of his extensive experience with assessment in statistics education. The colleague sent ideas for refining the items and suggestions for additional items that might help elicit students' conceptions of the discipline, all of which were incorporated in the final survey. For example, the colleague suggested an item regarding why statistics are needed, and so the item "How do you recognize if statistical methods are useful?" was added.

The survey items were then sent to another statistics education expert colleague from a third institution in a third region of the United States. The items were sent to this colleague because of her extensive experience and expertise with survey item-writing. This colleague sent feedback based on principles of item writing, such as breaking up items that contained two constructs. Specifically, the item that read, "What distinguishes statistics from math? ... from other sciences?" was broken up into two separate items. Also, based on feedback from the colleague, items were refined to have more focus. For example, the item that originally read, "What are the typical things a statistician does in their work each day?" was adjusted to read as, "What are 2–4 typical tasks a statistician might do in their work?" The adjustment, specifically exchanging the term *tasks* to replace *things*, was designed to elicit more relevant responses and avoid trivial responses such as "sit at a desk."

After further reflection and consideration of feedback, the items were sent again to the first colleague, who provided more feedback to refine the survey. For example, he suggested rearranging the order of items so that the questions about mathematics and sciences would not influence how students interpreted other questions. The final survey is given in Appendix A. The survey was intentionally kept short to encourage participation and keep it feasible to be administered briefly during class. The final survey included seven open-ended items regarding students' conceptions of statistics and two multiple-choice items from the GOALS instrument (e.g., Garfield et al., 2012) to allow for potential study of relationships between students' conceptions and content knowledge. Data from the GOALS items are not included in this analysis, but are mentioned here to provide the full context of the data collection.

3.4. DATA COLLECTION

The survey was administered online using *Google Forms*. The collection was scheduled so all participants would have experienced (or nearly completed) one college-level introductory statistics course at the time of their participation. Students in the upper-level course (Group One) participated on the first day of class, in February of 2018. No incentives for participation were offered to Group One, and only one student opted out of the research (100% response rate, 96% useable; $n_1 = 26$).

Participants in Group Two were solicited to participate toward the end of their introductory biostatistics course, in April of 2018. These students were asked to complete the survey outside of class, and the instructor asked all students to participate so responses could be used to inform her teaching. Students were offered a small amount of extra credit in exchange for their participation, and the extra credit was awarded regardless of whether students gave additional consent for their responses to be used in the study. The response rate for Group Two was 100% ($n_2 = 18$). In total, survey responses were collected from 44 students.

3.5. DATA ANALYSIS

The two approaches adopted for this study, phenomenography and grounded theory, share the call for iterative data analysis processes that involve steps such as becoming familiar with the data, looking for emerging categories/codes, refining categories/codes, and making connections (Kinnunen & Simon, 2012). As such, the three researchers spent time familiarizing themselves with the data and then analyzed the results using iterative rounds of qualitative coding, two of which were conducted independently and a third coming to consensus about each decision.

As is consistent with grounded theory, throughout the analysis process we fought to resist preconceived knowledge (e.g., Nolas, 2011). This brings about the well-documented debate about when in the investigative process researchers should seek out and review previous literature (e.g., Dunne, 2011). The research team intentionally did not refresh or add to their knowledge about prior research on students' conceptions of statistics until after the coding was complete. After completing coding, previous research regarding students' conceptions of statistics was reviewed (e.g., Gordon, 2004, Reid & Petocz, 2002) and comparison with that research helped to shape this report (Nolas, 2011). Two members of the research team were familiar with literature regarding experts' statistical thinking (e.g., Rossman et al., 2006; Wild & Pfannkuch, 1999) prior to conducting this study.

Despite efforts to try to avoid preconceived knowledge, the research team acknowledged biases and sought to make them as explicit as possible. Journal writing occurred before and throughout the coding process to document expectations. Common expectations included emphasis on procedures and formulas, statistics as useful, statistics as a subset or branch of mathematics, and negative views of the discipline, perhaps based from previous experiences using formulas without understanding how the procedures work or what they mean. The research team met to discuss journal entries to refine our thinking about these expectations as potential biases. This collaborative and reflective process elicited more preconceived biases as the researchers identified with and compared each other's shared biases. When unexpected themes emerged, the research team engaged in the grounded-theory process of abduction, whereby we make sense of the results by seeking explanations that made surprising findings seem reasonable (Nolas, 2011). This took the form of discussions among the research team, some of which are summarized in this paper.

Coding procedures The analysis involved three rounds of coding the entire data set. The first round of coding used the *in vivo* method, in which exact words or phrases from participant responses are used to form codes (Saldaña, 2015). The in vivo method was chosen to help ground results deeply in the data. To moderate potential biases caused by certain responses occurring in proximity to each other, all three researchers coded responses in a different randomly assigned order. In other words, each researcher analyzed the student responses according to a particular and distinctive randomly assigned sequence based on random numbers generated by spreadsheet software.

The unit of analysis was a single response on a single item of the survey. Although we reviewed the responses on each item individually, the responses were viewed together by student (rather than by question) to enable a student's previous answers to help provide context for interpretation of the student's other responses.

Before meeting together, the three researchers independently coded the entire data set and then reviewed their codes and grouped them into themes to identify 10–20 preliminary codes each. The research team met to discuss similarities and differences in independently-identified preliminary codes, and found similarities in these codes to establish a set of 24 overall *concept codes* (Saldaña, 2015) used for the second round of independent coding.

In the second round of independent coding the researchers reviewed all the data and coded each response according to the 24 agreed-upon concept codes. On an individual basis the researchers also added new candidate concept codes as they emerged from the data and seemed potentially promising. In this round of coding, the researchers maintained their same distinctive randomly assigned sequences of responses that were used in Round 1 so as to continue to moderate bias that may be caused by viewing certain responses in proximity to each other. The research group then met and discussed any additional concept codes that had been added and discussed whether it was appropriate to keep the concept codes or whether the codes were not as relevant as initially thought.

In what became a third round of coding, the research team together examined each student response and came to consensus on which of the concept codes applied to it. Multiple codes could be assigned to any one response. Responses were considered in a randomized order different from any of the individual researchers' individually assigned orders.

The process of coming to consensus about codes that applied to each of the responses served as a springboard for rigorous debate about the codes themselves. The debate included consideration of where to draw boundaries for each content code, the relevance of each code, specification of what the code did and did not represent, and delineating what should (and should not) be considered indication of a response eliciting in the code. Some codes were merged. After much discussion of the distinctions and whether they were relevant, notions of patience, perseverance, and persistence were combined into one code. Other codes were kept distinct after consideration of merging. For example, *Interpretation* and *Communication* were considered for merging because both have an element of helping others to understand ideas. After an iterative process of considering and re-considering whether they should be merged with each example of student response that arose, we eventually came to consensus that the two codes had nuanced differences important to keep distinct. *Interpretation* was used to indicate difficult concepts that needed unpacking; *Communication* was often used more generally to refer to an interpersonal skill.

When coding students' views as involving (merely) *Tests and Procedures*, a holistic evaluation of each student was made by looking at that student's responses to every question. More specifically, if a student referred to tests and procedures but later indicated a real-world application, they were not coded to have a *Tests and Procedures* view. The rest of the codes were assigned to each response separately. Finally, we merged our data so the codes were assigned to the students, not their responses. To be specific, the student was assigned the code if they included it for any of their responses.

To arrive at the final set of themes, the 27 codes that had survived to the third round of coding were considered for the extent to which they added depth and understanding to students' conceptions. The process was aligned with the two aforementioned criteria given by Marton and Booth (1997): (a) each category represents something distinctive, and (b) outcomes are parsimonious. For example, the *Robotic* code, which was designed to identify responses for which statistics was mostly computation-based, was removed based on criterion (a). Although represented by a large enough group of respondents, this code was not called a theme because there was not enough evidence to distinguish it from the *Tests and Procedures* theme, which had stronger evidence to support it. Regarding criterion (b), codes that were not strongly supported by the data (e.g., not represented by a critical mass of students) were dropped. Examples of such codes were *Data Visualization* and *Ideas*, the latter of which reflected notions of creativity and innovation. These themes simply were not represented by enough of the respondents to be justified in this study as themes, and in the interest of parsimony were dropped. In the end, a natural cut-point was 10%: the codes that emerged as themes were represented by at least 10% of participants.

After being immersed in the data, the research team began to discuss the most structurally substantive differences in student responses, as described for phenomenography by Marton (1986). The results of this discussion led the team to the final framework offered for describing students' conceptions of statistics.

Coding reliability testing To check the reliability of the coding process, a fourth independent statistics education researcher from another institution in a different region of the United States was solicited to use the codebook to code a sample of data. For an initial calibration exercise the new researcher coded a practice sample of 14 responses, each with 20 codes to be considered. Of this practice set of 280 decisions, 98% of the new researcher's decisions to code (or not) a particular response into a category agreed with the original determinations of the research group. After this calibration exercise, the independent researcher coded a larger random sample of 10 students' seven responses. These codes were compared with the research team's codes and 84% agreement was achieved. There were no major trends in the disagreement, however the worst-performing theme was *Tests and Procedures* (60% agreement), perhaps because of lack of clarity regarding the holistic approach taken to classify this theme. The overall percentage of agreement in coding was sufficient to consider the coding conducted by the researchers as sufficiently stable across time and people.

4. RESULTS AND DISCUSSION

Our theoretical framework guides the way the results are presented. We offer two sets of results: (1) as is common among phenomenographic and grounded-theory studies, the themes that emerged from the data (analogous to categories in phenomenography), and (2) as is aligned with phenomenographic approaches, a framework identifying the most structurally substantive differences that clarify how students communicated their conceptions of statistics. The framework emerged as the research team's primary focus of discussion after being immersed in the data (Marton, 1986). In this section we break from convention as we offer pertinent discussion of the themes along with the presentation of the themes. In particular, we offer discussion with respect to how students' conceptions reflect experts' conceptions, primarily based on Wild and Pfannkuch (1999) and Pfannkuch and Wild (2000). Broader findings, implications, and limitations are withheld and addressed in Section 5.

As suggested by Wild et al. (2018), we did not expect our novice students to have the same understanding as experts; we simply looked for similarities. Where there were potential glimmers of expert thinking, we identify these preliminary results as "seeds" of potential expert thinking. Meanwhile, here we would like to briefly acknowledge that all three of Makar and Rubin's (2009) aspects of informal statistical inference (generalizations, using data, uncertain language) were represented fairly well by student responses. It makes sense that students might better reflect the more informal aspects of statistics as they are still quite novice in their experiences of the discipline. As our primary research is concerned with comparing to experts' thinking, we proceed.

4.1. THEMES IN STUDENTS' CONCEPTIONS OF STATISTICS AND STATISTICIANS, AND CONNECTIONS TO EXPERT THINKING

The results in this section are presented as themes, not categories, so as to indicate that one response may reflect several themes. To illustrate this, throughout this section we will return several times to one particular student's response to Question 5. The student, whom we shall refer to as Adrian, had nothing otherwise particularly notable except that his is an example of how one response can illustrate several themes. Otherwise, naming of students is reserved for Section 4.2, when it is more helpful to point out themes across multiple responses offered by one student. Pseudonyms are used to maintain anonymity of all participants.

The twenty themes are given in Appendix B, along with descriptions and example responses for each. Not all themes offered are particularly relevant to our research questions (see, for example, the *Buzz Words* theme), but we have included them in the Appendix to support transparency (e.g., Shenton, 2004). Here we offer results and discussion of the themes most relevant to our research. Throughout this section, the use of italics within quotations indicates emphases that have been added by the authors so as to point out the most relevant features of a response.

Variation and Uncertainty theme Variation and uncertainty are central to statistics (e.g., Davidian & Louis, 2012). Although the exact term *variation* was not common among student responses, about 40% of students elicited nuances of uncertainty. Two students used the term *uncertain* explicitly. One of these students made several references to uncertainty, noting that a statistical method "may be useful if one can make predictions with statistical significance and explain *uncertainty*." Later, the student said statistics differs from mathematics because it is "more focused on *uncertainty* and making *predictions.*" Comparisons to the sciences revealed this conception as well (e.g., "In these [*sic*] sciences, things can be observed, but statistics often study *uncertain* outcomes"). For the other student, an explicit reference to uncertainty was revealed when students were asked to compare statistics to the other natural sciences: "Statistics is a lot more *uncertain* than typical math."

Aside from these two students, students who alluded to uncertainty did so only indirectly, using subtle terms carrying nuances of uncertainty. Terms such as *predict, chance, likelihood,* and *probability* were common. Adrian wrote, "Statistics is a specific category of math, where there is still theory behind it, but it is more used to *predict* what will happen under certain circumstances and how *likely* different out comes are." Other examples include phrases such as: a statistician will "find *probabilities*"; statistics is "the study of the *likelihood* of events…"; statistics is "the study of … what differs from *chance*."

A few students articulated uncertainty by suggesting that the results of analyses may involve a range of possible solutions (e.g., statistics is "a science of calculating probability and giving answers with *probable range*"; "Math is finding solutions to equations and statistics is finding a *numeric range of solutions* that helps give answers to everyday and complex questions"). One student articulated this very clearly as the idea that "statistics is predicated on the notion that we *cannot figure out an exact answer* (as we would if we were doing calculus and solving for a variable) but instead that we seek to know *only approximately where the correct answer may be*."

The closest that many students came to mentioning variation (and we did not include these instances in the 40% figure given above) was via mention of variables. Students seemed to refer to variables as attributes studied, without emphasizing their inherent variation. For example, one student said it was important for statisticians to be "able to understand how variables affect results..." Although *vary* is the root of the term *variable*, we did not interpret this response to hold elements of variation; it seemed the student was referring more to observations rather than things that vary. Other students approached notions of variation by referring to differences. For example, one student commented that a statistician may "run a test for a difference in means." The research team considered these references not explicit enough to qualify as a mention of variation and uncertainty.

Whereas Pfannkuch and Wild (2000) note that consideration of variation is foundational to experts' notions of statistical thinking, the students' mention of variation tended to underrepresent this central facet of statistics. In fact, none of the students used the terms *variation, variability,* or *varies* in any of the open-ended responses. The closest that many students came was to refer to *variables*, but in the contexts that this term was used it generally seemed as the term was used more to represent a predictor, or an observed aspect of cases, without real note of the variability.

Notions of uncertainty, which was much more prominent, may serve as seeds of variation. Seeds of variation may also include references to *differences* (e.g., different treatments, group differences), and references to change (how one thing changes due to another).

Logic and Mathematics theme One of the most prevalent themes included references to logic and mathematical skills required for statistical work. Some students elicited this theme in response to survey Question 3, offering that statisticians need to have logical thinking or mathematical abilities to be successful. Many of the students whose responses indicated a mathematical theme did so by suggestion on Question 5 that statistics was a subset of math. Thirty-four percent of participants indicated a subset relationship, despite the fact that the (leading) survey question asked, "What distinguishes statistics from math?" Still, responses such as Adrian's ("statistics is a specific category of math…"), and others ("statistics is a type of math," and "statistics is the math used to analyze data in the natural sciences") indicated that statistics is a subset of mathematics.

Many other students did not specify that statistics is a type of math, but indicated statistics *uses* math or requires math and logic skills. Some responses were difficult to classify (e.g., statistics is "the mathematical study of risk and analyzing of data," "statistics is the back bone of math that seems most interesting to most people in math…"); the former was *not* coded by the research team to imply subset but the latter was.

Another relationship students offered between statistics and mathematics posited statistics as an application of mathematics. For example, returning to Adrian's response to Question 6, "Statistics is a specific category of math, where there is still theory behind it, but it is more used to predict what will happen under certain circumstances and how likely different out comes[*sic*] are." Adrian seemed to consider application to be one of the biggest differences between mathematics and statistics.

Although Wild and Pfannkuch's experts did not explicitly mention mathematical skills as a prerequisite for expert statistical thinking, the skills are implied by some of the other aspects of their framework such as transnumeration, considering variation, and working with statistical models: all of which require sophisticated mathematical skills. Moreover, in their model, logic is listed explicitly as a type of thinking that statistical experts need. Statistical experts, however, may take issue with the relationships that students offered between statistics and mathematics. Many students identified statistics as a subset or application of mathematics, whereas Rossman et al. (2006), for example, identify several important distinctions between the disciplines, including the different crucial role of context and the central role of the data collection process in guiding appropriate conclusions.

Tests-and-Procedures and Real-World Problem-Solving themes The other most-prevalent theme (represented by 66% of participants) was that statistics involves solving real-world problems. This theme included references to specific real-world applications (e.g., "determining the effect of well water on people near a landfill") as well as general references to solving problems in applied settings (e.g., in response to Question 4: "When they can use it in everyday life"). Occasionally we saw evidence of this theme when participants compared statistics to mathematics or other sciences (e.g., "…statistics is more about finding meaning in the data and numbers whereas math is just about using equations to find a number/answer").

The *Real-World Problem-Solving* theme seemed to contrast with another strong theme (represented by 43% of participants), regarding statistics as the use of tests and procedures. Responses coded as the *Tests and Procedures* theme tended to offer a slurry of robotic procedures (e.g., "Collect data, determine whether results were significant."). In contrast, a *Real-World Problem-Solving* conception was coded for responses such as, "(Statistics) allows people to receive evidence to questions about *every day things*. It requires data collection and analysis and a inquisitive mind about the *world around you*." As indicated by percentages that sum to more than 100%, we allowed student responses to be simultaneously classified as containing both of these seemingly contradictory themes. This is consistent with previous research suggesting conceptions of statistics may fall on a hierarchical spectrum: at one end a series of tests and procedures and at the other end a way of solving real-world problems and making meaning (e.g., Reid & Petocz, 2002; Gordon, 2004). As previous research was intentionally reviewed *after* conducting our analyses, it was striking that these results emerged independently (albeit one researcher recorded in her pre-study journaling that she anticipated a test-and-procedure conception).

Patterns theme: "Seeds" of causes Another theme represented by nearly half of participants was related to looking for patterns, trends, or associations in data. We defined this theme to include any reference to a relationship. For example, we turn to Adrian's response: "Statistics is a specific category of math, where there is still theory behind it, but it is more used to *predict* what will happen *under certain circumstances* and how likely different out comes[*sic*] are." After much deliberation, the research team understood their response to imply a pattern because Adrian says that under certain circumstances, statistics can be used to predict events. This implies a pattern, to a moderate extent, between circumstances and other outcomes. Other examples of statements eliciting this theme include references to trends and associations such as "analysis of data to understand its relationship," "find/make correlations," and "predict future outcomes."

Experts, as described by Wild and Pfannkuch (1999), are ever on a very important "quest for causes." That is, they are always looking to try to find the reason behind variation. About 10% of our participants explicitly referred to causes or causal claims. A more common and related theme in our data was a quest for *patterns*, *trends*, and *relationships*, which we saw in 48% of participants' responses.)

We recognize there is a difference between causes and patterns (correlation is not causation!). However, Wild and Pfannkuch (1999) describe the first step of investigating causes as looking for patterns in variables. Some patterns may indicate an underlying signal in the data that can be used to predict or explain variation in the data. The rest of the variation is interpreted as random and does not help explain variation in the data. A central idea in statistical modeling is making reasonable distinctions between variation from patterns and variation from random phenomena.

In some cases, the patterns or associations between the variables can be attributed to causal relationships between the known variables in a study. In other cases, the pattern is the result of relationships between another variable that is not included in the study but is the ultimate source behind the observed pattern. In either case, finding patterns between variables is a first step in identifying causal relationships. In other words, for causation to exist, there must first be an association, or pattern, between two or more variables. As association is a pre-requisite for causation, we imagine that reasoning about notions of pattern may be a springboard from which novices can move toward the next step of investigating potential causes for the pattern.

Persistence theme One of the themes that was least reflected in our preconceived ideas and emerged from the data was the extent to which students communicated that statistics required persistence, hard work, and other terms related to perseverance beyond what is natural or easy. Question 3 (“List 2-4 character traits a successful statistician needs to have”) revealed this aspect of students’ conceptions most often. Responses included notions of mental endurance, perseverance, persistence, patience, resilience, dedication, and hard work, all of which we grouped under “persistence.” This theme was evident from 40% of participants.

That students identified statistics as requiring perseverance aligns with many previous studies that suggest students think statistics is difficult (e.g., Malaspina, 2018). Yet difficulty is often viewed as an attitude toward statistics rather than a characteristic of the discipline itself. The survey questions were designed to bring about students’ *conceptions* of statistics rather than their *attitudes* toward it. This may be an indication of the strength of this particular attitude.

Experts also acknowledge that statistics requires perseverance. Wild and Pfannkuch (1999) explicitly identify perseverance as a disposition of the discipline, and add that experts find engaging, interesting problems to be helpful in supporting perseverance. The authors posit that students need their teachers to choose engaging problems to help them persevere through difficult tasks in their statistics courses.

Collect Data, Analyze Data, and Conclude themes Two more themes were elicited by over half of participants, and were often found together. One involved references to collecting data, and the other references to analyzing the data. The last three steps in the PPDAC cycle (Data, Analysis, Conclude) were strongly represented in responses (57%, 59%, and 30% of participants, respectively). In many cases, particularly in response to Question 1 (see Appendix A) regarding statisticians’ typical tasks, students used close derivations of the terms “*collect*” or “*analyze*.” For example, one student responded, “Gather data, analyze/look for correlations in data.” Another student said, “Collect data, run tests to analyze the data, use the data to make assumptions about a specific population.”

The earlier steps in the PPDAC (Problem and Plan) were much less common. One of the rare exceptions was a student who referred to the *Problem* phase, stating a statistician’s typical tasks are to “first, identify a question or problem...” For most students, the closest students got to mentioning the Problem phase of the PPDAC cycle was to refer to the fact that statistics could be used to solve real-world problems. We did not see much evidence of the Plan phase, except for a few references to planning data collection, which might be better categorized in the Data phase (data collection).

This emphasis on the DAC steps of the cycle is consistent with Wild et al. (2018), who suggest that naïve students immediately try to find a suitable technique (Data, Analysis steps of PPDAC), whereas the experts began by identifying the scientific question (Problem step of PPDAC). Moreover, the students’ lack of inclusion of the first two steps of the PPDAC cycle suggests that respondents had similar conceptions to the *clients* described by experts in the studies by Wild and Pfannkuch (1999); they did not recognize that statisticians can (and ought to) play an important role in the problem and planning phases. As noted by an expert in Pfannkuch and Wild (2000), “Getting from the existence of a problem to questions to be answered by gathering data constitutes a very substantial part of the work" (p. 140).

Our survey questions did not elicit much from our students in terms of recognizing the need for data as opposed to personal experiences and anecdotes. This is in contrast to results of Pfannkuch and Wild (2000) who found that a foundation of expert thinking (Dimension 2) is recognition of the need for data. No participants mentioned the inadequacy of anecdotes.

Responses that articulated the role of data explicitly were elicited by Question 5, which asked students to identify the difference between statistics and mathematics. Responses included "The use of *data* and conclusions found based on *data*"; "Its emphasis on *data* and the purpose for which numbers are collected"; and "Statistics is about interpreting the *data* whereas math is about computing numbers and using theories." One student summed it up in response to Question 4, saying, "If there is data involved, there is probably great use for statistics." In short, students tended to turn to data as a way to distinguish statistics and mathematics.

Although students did not often explicitly express the need for data, student responses provide some preliminary evidence that our students consider data as central to statistics. As suggested by the word cloud in Figure 2, the term *data* was used in responses far more than any other term (157 times). For comparison, the second most frequent term (statistics) was used only 77 times, and no other words were nearly as frequently represented. (The third highest term had only 34 occurrences.) The word cloud provides a visual depiction of how the term *data* played a strong role in student responses.

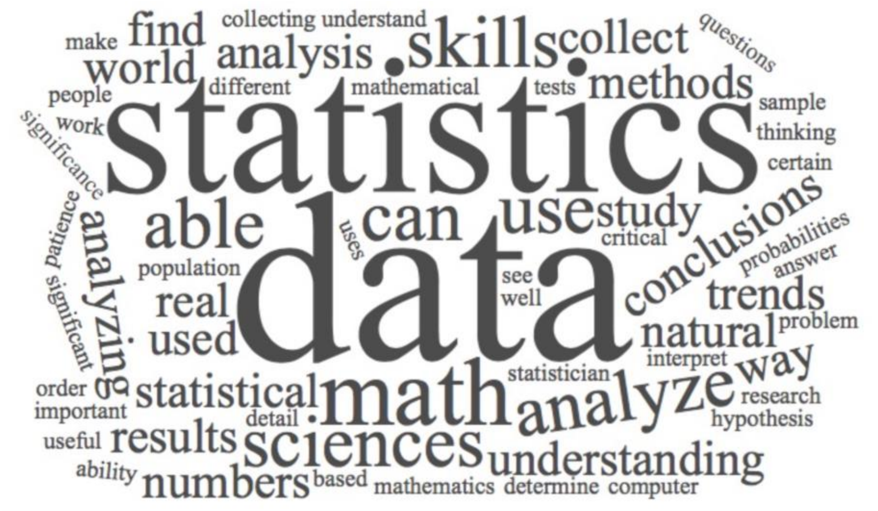

Figure 2. Word cloud of terms appearing 8 or more times in open-ended responses to survey Questions 1–7. Size of word is roughly proportional to the square root of usage. (Created on www.WordClouds.com, Zygomatic, 2018)

Although not explicitly articulated as important, the students seemed to have an underlying sense that data are central to the discipline. In short, although we did not see evidence that students describe statisticians as having an impulse to collect data as voiced by the experts, the students still seemed to generally have a sense that data are central to the discipline.

Communication theme The need for good communication skills was noted by many (40%) participants. The emphasis, however, was much less focused on communicating to a particular client: in our data, communication skills were expressed in general terms. For example, when asked about skills needed for statisticians one student included, "…being able to communicate and work well with others." It is interesting that 25% of the students who indicated communication skills also mentioned some element of people skills in tandem with communication. In fact, only three students mentioned people skills without also referring to communication. This suggests that many students may recognize that these two skills often occur in tandem.

Communication played an important role in Wild and Pfannkuch's records of experts' interactions with clients (2000). They note that experts work very hard at "translating into the language of the client" (p. 150). Wild and Pfannkuch (1999) also suggest communication skills are required for extracting vital information from clients. In our data, the focus was more in the other direction, as a way of communicating results. This difference may be another reflection of the lack of the earlier steps in the PPDAC cycle.

Ethics and Integrity theme For a few students (n = 8, 18%), statistics involved ethical decisions. For example, one student commented that statistics is different from other natural sciences in that "statistics requires a great amount of integrity because data can be manipulated and conveyed in many different ways unlike other natural sciences." Other students alluded to this theme by referring to bias, suggesting, for example, that statisticians need to avoid bias when collecting, analyzing, or reporting data-based results.

Wild and Pfannkuch (2000) document some ways that ethics plays a role in experts' experiences conducting statistical consulting, whether there be financial incentives to adjust results or whether it is just difficult to give results that are not aligned with what the client wanted. The ethical issues described by the experts had slightly stronger emphasis on outside influences, particularly clients.

4.2. A MULTI-DIMENSIONAL FRAMEWORK FOR STUDENTS' CONCEPTIONS OF STATISTICS

Our results corroborate with previous studies of students' conceptions of statistics that suggest students' conceptions vary by the extent to which they view statistics as a series of tests and procedures versus a tool to solve real-world problems (Bond et al., 2012; Gordon, 2004; Reid & Petocz, 2002). We found preliminary evidence, however, to suggest that there is room to expand upon that framework by adding other dimensions. Most notably, our research team focused on the extent to which students recognize variation and acknowledge the uncertainty of statistical conclusions (e.g., Rossman et al., 2006). We found evidence to suggest that these aspects of students' conceptions play a distinctive part in illustrating students' concepts of statistics.

Figure 3 illustrates the beginnings of a framework we propose. Whereas previous research (e.g., Bond et al., 2012; Gordon, 2004) categorized students along one spectrum (displayed horizontally), our research suggests a new dimension (displayed vertically) representing the extent to which students acknowledge uncertainty. Although the image displays these themes as orthogonal, this may not

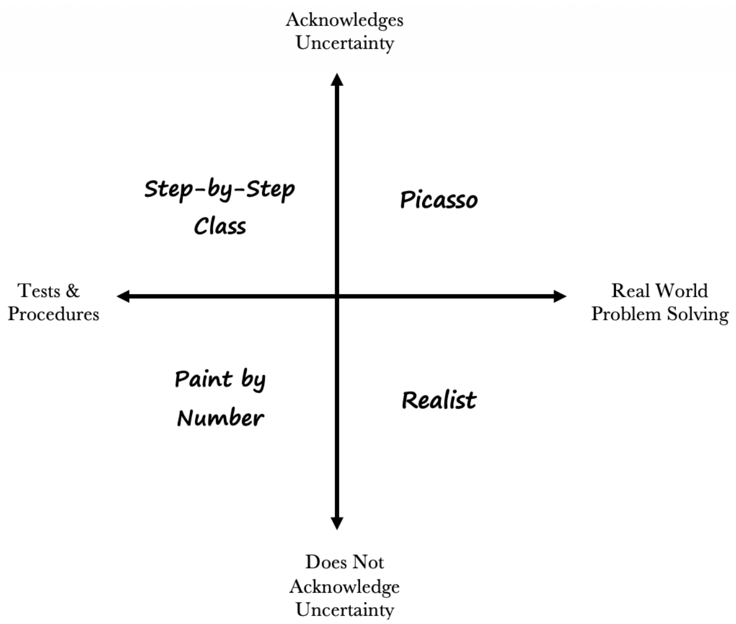

Figure 3. Diagram illustrating a framework for students' conceptions of statistics

necessarily be the case. Still more dimensions could be added. As is aligned with our phenomenographical approach, however, we focus on two that were the subject of the discussion among the research team as the most structurally substantive in categorizing students' conceptions (Marton, 1986).

We describe in detail some possible conceptions using this framework through the metaphor of different styles of painting to illustrate the differences between conceptions. For each quadrant we will provide an example student who offered responses reflecting this conception and a painting style to serve as a metaphor for that conception. As is aligned with phenomenography, we do not assume the students' conceptions are stable over time; these students are reflecting the given conceptions—which are stable as a collective entity. Each student is assumed to reflect the conception in the particular context and moment of the study at hand.

Paint-by-number: Less recognition of uncertainty with the idea that statistics is a collection of procedures A student in the bottom left quadrant sees statistics mostly as a series of tests and procedures conducted to arrive at conclusive results. This student may think of statistics as completing a paint-by-number activity (see for example, Figure 4). In paint-by-number activities the colors and picture are predetermined, and the artist's job is to follow the instructions filling in the colors corresponding to the indicated number. The objects of the art are often lovely, albeit trivial, and the images typically do not call for much deeper interpretation.

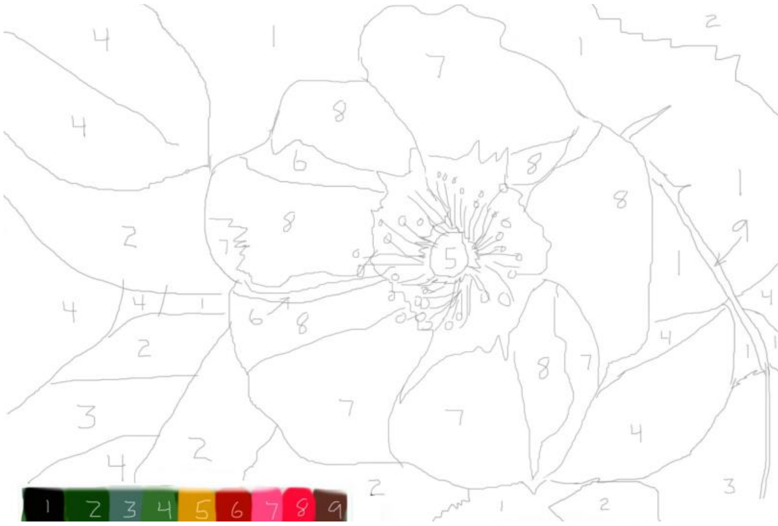

Figure 4. An example of a paint-by-number activity; there is not much meaning-making or uncertainty about the image. Reprinted under Creative Commons 2.0 license (https://creativecommons.org/licenses/by/2.0/deed.en) with attribution: J. Blodget (2009). *Paint by number*. [Brushes on an iTouch with a Pogo stylus].

This conception may be demonstrated by a student whom we will call Eloise. Her survey responses were

Q1: Typical tasks include evaluating data and interpreting graphs.
Q2: Skills that a successful statistician needs is to be good in math and to be able to look at data and find trends.
Q3: Character traits include neat, organized, and hardworking.
Q4: If they are efficient when used and lead you to the right answer.
Q5: Statistics deals more with analyzing numbers and being able to interpret them for any trends.
Q6: They deal less with physical components and more to do with numbers and data.
Q7: Statistics is the practice of collecting and analyzing numerical data in order to find any trends or patterns.

Here Eloise did not put much emphasis on making meaning of the world or solving real-world problems; her conceptions seem to be limited to executing procedures. At the same time, she refers to "the right answer" and does not allude to many elements of variation or certainty in results.

Step-by-step instruction: More recognition of uncertainty with the idea that statistics is a collection of procedures A student in the top left quadrant might view statistics as a set of tests and procedures with results that are not certain. These conceptions may be likened to artwork created in a novice painting class where the day's concept and colors are predetermined by the instructor. The students—who have little choice in the subject—follow the instructions of the instructor as best as they can, and still they may not arrive at an image that is entirely discernible. Students with analogous conceptions of statistics may follow the steps (tests and procedures) yet still recognize the results are not exact. Examples of paintings representing this conception of statistics are given in Figure 5.

Figure 5. Two paintings from the same painting class, illustrating step-by-step artwork that still contains uncertainty about the original model. Left: E. Fry, 2013, *Moonlight Tree*. [Acrylic on canvas]. Printed with permission. Right: K. Edwards, 2013, *Here Comes the Moonlight*. [Acrylic on canvas]. Printed with permission.

Our best example of a student offering conceptions in this quadrant may be Henry, whose responses were

Q1: Hypothesis Testing, Computing Probability, Typing up Data, Running error tests on statistical software
Q2: Computer skills, calculus
Q3: Patience, non-biased attitude, eye for detail
Q4: Use tests to see if the results are significant, repeat the test over the years with different groups or larger populations.
Q5: While math is used for specific answers, statistics is used to make approximations, connections, and inferences.
Q6: Statistics can be applied to many kinds of sciences and cannot be directly observed but must come from a compilation of tests.
Q7: The study of the likelihood of events based on real life data.

Henry's responses indicate uncertainty via phrases such as "statistics is used to make approximations," "error tests," "likelihoods," and "repeating the test with different groups or larger populations." Meanwhile, the emphasis is less on solving real-world problems and more on computing the actual tests. Henry does, admittedly, mention that statistics can be applied to sciences and mentions real-life data, so his conception may belong in the top middle of the quadrant rather than top left.

To illustrate the importance of the vertical dimension in Figure 3 and underline our argument for adding this dimension, we take a moment to compare the responses from Henry and Eloise. Both students describe a conception that falls on the left side of Figure 3, as both communicate a lot of procedural aspects of statistics without much emphasis on real-world applications or making meaning. Henry, however, uses descriptions that include much more emphases on potential errors, approximations, and inferences. Henry seems to illustrate a conception that is more aligned with experts' in the sense that he recognizes the role of variation and uncertainty in the discipline. In comparison, Eloise's responses indicate making patterns and finding "right answers," which is much less aligned with experts' conceptions. By including the vertical dimension in the framework, we are able to acknowledge this distinction.

The Realist: Less recognition of uncertainty with the idea that statistics can solve real-world problems Conceptions in the bottom right quadrant of the diagram portray statistics as a way of solving problems or making meaning, while holding results conclusively. This conception may be analogous to realist painting. As with the Rembrandt self-portrait (Figure 6), the art contributes to our conception of what the artist looks like, but does not require much interpretation. Students in this category may capture an image or a problem and seek to answer it, but they consider their results without much uncertainty.

Figure 6. An example of a real-world image that does not require much interpretation under uncertainty. Reprinted under Creative Commons 1.0 License (https://creativecommons.org/licenses/by/1.0/) with attribution: R. van Rjin (1632–1636). *Rembrandt wearing a soft cap: full face: head only*. [Etching on paper]. Statens Museum for Kunst.

This conception may be demonstrated by Bess, whose responses were

Q1: Organizing and cleaning data, developing trials, testing data for correlations
Q2: Problem solving, coding, intuition, mental endurance
Q3: Communicative, teamwork, persistent
Q4: Any study they see that brings new light upon a subject gives recognition that statistics is useful.
Q5: Statistics is about solving real world problems, math can be focused more on theoretical aspects.
Q6: Statistics can and is used in natural sciences but the pure subject deals with statistics as an entire field (theory, etc.) while other sciences use statistical methods within the context of their principles.
Q7: A mathematical science meant to find significant correlations in real world data.

Bess explicitly states that statistics is about "solving real-world problems" and finding "significant correlations in real-world data." Yet her responses do not illustrate much recognition of variation or uncertainty. One might be able to argue that her use of "significant" may be drawing from notions of random chance, however, Bess may also be using the term in a less technical sense, as a synonym for "meaningful."

Picasso: More recognition of uncertainty with the idea that statistics can solve real-world problems A student in the top right quadrant is most aligned with our conception of statistics and that of experts; that is, they view statistics as a way of solving real-world problems but holding their results with uncertainty. We liken these students' conceptions loosely as similar to many of Picasso's paintings (e.g., Figure 7); there is a lot of meaning behind the final product, but also a lot of uncertainty about the original subject; geometric forms and disorganized facial features are obvious clues that one should not assume the image looks exactly like the model!

Figure 7. A meaning-making image with uncertainty (the precise appearance of the model is indiscernible). Reprinted under Creative Commons 2.0 license (https://creativecommons.org/licenses/by/2.0/) with attribution: P. Picasso, 1935, *Head of a woman*. [Oil on Canvas]. Sotheby's, London.

An example of a student whose conceptions may fall in this quadrant is Zephaniah:

Q1: Analyze sports performance, analyze financial performance, create econometric models to explain a variable and make predictions, analyze risk
Q2: Critical thinking ability, problem solving skills, computer program skills
Q3: Detail oriented, integrity
Q4: A method might be useful if one can make predictions with statistical significance and explain uncertainty.
Q5: Statistics are more focused on uncertainty and making predictions.
Q6: In these sciences, things can be observed, but statistics often study uncertain outcomes.
Q7: The mathematical study of risk and analyzing of data.

Zephaniah offers many examples of real-world contexts, and articulates a recognition of uncertainty explicitly (e.g., "Statistics are more focused on uncertainty and making predictions"; "statistics often study uncertain outcomes") and implicitly (e.g., "risk," "predictions").

Here it is appropriate to pause and underscore the fundamentally different conceptions of Bess and Zephaniah. Both students indicated that statistics involves solving problems in the world outside of academics. However, there is a distinctive difference in the extent to which they communicate the uncertainty with which they hold their results. As indicated above, Zephaniah alludes to uncertainty explicitly and implicitly throughout his response. In contrast, Bess' response does not have nearly as much explicit acknowledgement of variation or the central role that uncertainty plays in the discipline.

5. CONCLUSION

This study was conducted to explore students' conceptions of statistics and compare them to experts' conceptions of statistics. We gathered responses to open-ended survey questions from 44 undergraduate students who had completed (or nearly completed) an introductory statistics course. Based on the themes developed via several rounds of qualitative coding, we offered themes in students' conceptions of statistics. The themes of *Tests-and-Procedures* versus *Real-World Problem-Solving* align with previous research on students' conceptions of statistics (Bond et al., 2012; Gordon, 2004; Reid & Petcoz, 2002). Our results align with research conducted by Findley and Kaplan (2018), who explored how mathematical thinking plays a role in conceptions of statistics. In this paper we also suggest new themes, and of particular interest is the extent to which student conceptions acknowledge variation and uncertainty. We used a metaphor of different painting styles to illustrate a framework for characterizing students' conceptions.

We also examined how our participants' conceptions compare to experts' conceptions of statistics. Similarities include notions of ethics, communication, and the Collect Data, Analyze, and Conclude parts of the PPDAC cycle (Wild & Pfannkuch, 1999). Some areas in which our students were lacking were the Problem and Plan stages of the cycle. We also offered a few "seeds" that may be cultivated to reflect more expert-like thinking, such as the *Patterns* theme as a seed for "the quest for cause" (Wild & Pfannkuch, 1999), and the *Differences* theme as a seed of consideration of variation.

In this section we conclude with a discussion of limitations, implications, and areas for future research.

5.1. LIMITATIONS

Despite efforts to identify and isolate preconceived ideas, the researchers brought their own expectations and biases when interpreting student responses. This was particularly true of the Tests-and-Procedures dimension; even without reading the prior literature some of us had prior expectations that this might be revealed by the data. The results of this study produced strong support for this dimension, but the conclusion must be tempered by preconceived expectations that the dimension might exist.

Another limitation arises from the fact that the data analysis was based on short-answer survey responses, not in-depth interviews. We temper this critique with the nature of the goals of this study: to get a sense of students' conceptions of the discipline in a non-pressure setting, and to identify possible aspects of variation between student conceptions. We acknowledge that follow-up interviews will allow deeper probing into what students meant by particular phrases.

As the subjects were not a random sample, the authors do not attempt to generalize these results to broader populations or claim that the results are representative of all students who have taken an introductory undergraduate statistics course. Instead, the results and methods are reported with transparency to encourage reproducibility and trustworthiness, so that the reader may determine the transferability of this study to other student groups or regions (Guba, 1981; Shenton, 2004).

5.2. IMPLICATIONS FOR TEACHING

Education research charges teachers with the challenge of understanding students' conceptions and building on them rather than considering their students as empty slates to be filled with new knowledge (e.g., Bransford et al., 2004). This research offers teachers a variety of conceptions about statistics that students may hold. We also offer suggestions as to how some conceptions may serve as seeds of expert thinking that may be cultivated to help students build more expert-like conceptions of statistics. For example, students' references to *differences* within variables may be an opportunity to help students develop a stronger sense of the central role that variation plays in statistics. As another example, exploring *patterns* that students notice may help them begin to reflect Wild and Pfannkuch's (1999) expert-like "quest for cause."

Current recommendations in statistics education include teaching statistical thinking, and teaching statistics as an investigative process which includes problem-solving and decision making in a real-world context (GAISE College Report ASA Revision Committee, 2016). Of the five steps in the

PPDAC cycle (Wild & Pfannkuch, 1999), the last three (Data, Analysis, and Conclude) were strongly represented in students' responses, whereas there was not much mention of the first two (Problem, Plan). Instructors may wish to examine whether their curricula underemphasize identifying problems and planning. It may be valuable to supplement with activities and assessments that include posing research questions and describing how a data collection process could be conducted.

5.3. IMPLICATIONS FOR FUTURE RESEARCH

In this research we set out to explore students' conceptions of statistics and to compare them to experts' conceptions. We identified several candidate dimensions of students' conceptions, including the extent to which they perceive uncertainty and variation to play a role in the discipline. This, along with several other dimensions proposed, sets the groundwork for future research, and there are many opportunities to build on these results. Follow-up studies could examine the extent to which the themes emerge for other groups of students in other contexts, and interviews could be conducted to explore whether the identified candidate themes are relevant, related, and/or distinct from each other.

Perhaps the most pertinent future research question would examine how to grow some of the seeds of students' conceptions of statistics into conceptions that are more aligned with those of experts. Here we identified potential starting points, but questions remain as to how to proceed in developing these into mature conceptions. Most studies conducted thus far on students' conceptions are cross-sectional (e.g., Reid & Petocz, 2002) or for short periods of time. Future work could gather rich interview data, revisiting interview subjects longitudinally—perhaps even beyond their degree programs and over the course of their careers. Such studies may examine the extent to which the dimensions offered here are relevant, explore whether there are relationships between dimensions, and identify learning trajectories for students' views of statistics. Moreover, such interviews might probe deeper into the meaning of the terms as used by students (e.g., Bess's use of the term "significant"). Whereas some studies (e.g., Dunn et al., 2016; Kaplan et al., 2009) have been conducted on lexically ambiguous words (e.g., *mean*, *significant*), here we call for study of words and phrases that are ambiguous in terms of what students mean when *they* use these phrases, which in turn would help illustrate their conceptions.

Finally, relationships could be explored between students' conceptions of statistics and their success in the discipline. Success in the discipline could go beyond passing courses or scoring well on quality assessments: do students with certain conceptions tend to take more statistics courses, pursue majors and minors in statistics, or pursue graduate study? Do students with certain conceptions tend to avoid further statistical study? Studies of these relationships could help inform recruitment of students into the discipline and preparation of a workforce that can attend to the data deluge.

ACKNOWLEDGEMENTS

This work was supported by the Natural Sciences Summer Undergraduate Research Program at Pacific Lutheran University. We gratefully acknowledge the contributions of Matthew Beckman for his input on the survey items and Katherine Edwards for sharing her artwork.

REFERENCES

Blodget, J. (2009). *Paint by number.* [Brushes on an iTouch with a Pogo stylus]. Licensed under Creative Commons 2.0. https://creativecommons.org/licenses/by/2.0/
[Online: https://www.flickr.com/photos/jimblodget/3868581531/]

Bond, M. E., Perkins, S. N., & Ramirez, C. (2012). Students' perceptions of statistics: An exploration of attitudes, conceptualizations, and content knowledge of statistics. *Statistics Education Research Journal*, *11*(2), 6–25.
[Online: https://iase-web.org/documents/SERJ/SERJ11(2)_Bond.pdf]

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2004). *How people learn* (Expanded ed.). National Academy Press.

Chance, B., & Rossman, A. (2017). *Investigating statistical concepts, applications, and methods.* [Online: http://www.rossmanchance.com/iscam3/]

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

Creswell, J. W. (2007). *Qualitative inquiry and research design* (2nd Ed.). SAGE Publications.

Davidian, M., & Louis, T. A. (2012). Editorial: Why statistics? *Science*, *336*, 12.
[Online: https://science.sciencemag.org/content/336/6077/12.full]

delMas, R. C., Garfield, J., Ooms, A., & Chance, B. L. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, *6*(2), 28–58.
[Online: https://iase-web.org/documents/SERJ/SERJ6(2)_delMas.pdf]

Dunn, P. K., Carey, M. D., Richardson, A. M., & McDonald, C. (2016). Learning the language of statistics: Challenges and teaching approaches. *Statistics Education Research Journal*, *15*(1), 8–27.
[Online: https://iase-web.org/documents/SERJ/SERJ15(1)_Dunn.pdf]

Dunne, C. (2011). The place of the literature review in grounded theory research. *International Journal of Social Research Methodology*, *14*(2), 111–124.
[Online: https://www.tandfonline.com/doi/abs/10.1080/13645579.2010.494930]

Edwards, K. (2013). *Here comes the moonlight*. [Acrylic on canvas].

Findley, K., & Kaplan, J. (2018). What is statistics? Examining the disciplinary beliefs of statistics teaching assistants. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS-10, July, 2018*), Kyoto, Japan. International Statistical Institute.
[Online: https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_C175.pdf?1531364321]

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: A pre-K–12 curriculum framework*. American Statistical Association.
[Online: www.amstat.org/Education/gaise/]

Fry, E. B. (2017). Students' conceptual understanding of the relationship between study design and conclusions in an introductory statistics course. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS-10, July, 2018)*, Kyoto, Japan. International Statistical Institute.
[Online: https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_8C3.pdf?1531364292]

Fry, E. (2013). *Moonlight tree*. [Acrylic on canvas].

GAISE College Report ASA Revision Committee. (2016). *The GAISE (Guidelines for Assessment and Instruction in Statistical Education) College Report*. American Statistical Association.
[Online: www.amstat.org/Education/gaise/]

Garfield, J., delMas, R., & Zieffler, A. (2012). Developing statistical modelers and thinkers in an introductory, tertiary-level statistics course. ZDM, *44*(7), 883–898.

Gordon, S. (2004). Understanding students' experiences of statistics in a service course. *Statistics Education Research Journal*, *3*(1), 40–59.
[Online: https://iase-web.org/documents/SERJ/SERJ3(1)_gordon.pdf]

Guba, E. G. (1981). Annual review paper: Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology: A Journal of Theory, Research and Development*, *29*(2), 75–91.

Hedges, S., & Harkness, S. S. (2017). Is GAISE evident? College students' perceptions of statistics classes as "almost not math." *Statistics Education Research Journal*, *17*(1), 337–356.
[Online: https://iase-web.org/documents/SERJ/SERJ16(1)_Hedges2.pdf]

Holcomb, J. (2015). Consider a career in statistics. *Chance*, *28*(4), 30–33.
[Online: www.tandfonline.com/doi/abs/10.1080/09332480.2015.1120125?journalCode=ucha20]

Kaplan, J. J., Fisher, D. G., & Rogness, N. T. (2009). Lexical ambiguity in statistics: What do students know about the words association, average, confidence, random and spread? *Journal of Statistics Education*, *17*(3), 1–22.
[Online: http://jse.amstat.org/v18n2/kaplan.pdf]

Kinnunen, P., & Simon, B. (2012). Phenomenography and grounded theory as research methods in computing education research field. *Computer Science Education*, *22*(2), 199–218.
[Online: https://doi.org/10.1080/08993408.2012.692928]

Lock, R. H., Lock, P. F., Morgan, K. L., Lock, E. F., & Lock, D. F. (2013). *Statistics: Unlocking the power of data*. Wiley.

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, *8*(1), 82–105.
[Online: https://iase-web.org/documents/SERJ/SERJ8(1)_Makar_Rubin.pdf]

Malaspina, M. M. (2018). Psychological factors associated with academic performance in statistics. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS-10, July, 2018)*, Kyoto, Japan. International Statistical Institute.
[Online: https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_5E3.pdf?1531364271]

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

Marton, F. (1981). Phenomenography—Describing conceptions of the world around us. *Instructional Science*, *10*(2), 177–200.
[Online: https://link.springer.com/article/10.1007/BF00132516]

Marton, F. (1986). Phenomenography—A research approach to investigating different understandings of reality. *Journal of Thought*, *21*, 28–49.
[Online: https://www.jstor.org/stable/42589189]

Marton, F., & Booth, S. (1997). *Learning and awareness*. Lawrence Erlbaum.

Marton, F. (1994). The idea of phenomenography. In R. Ballantyne & C. Bruce (Eds.), *Proceedings of Phenomenography: Philosophy and practice conference (November, 1994)*, Brisbane, Australia. Centre for Applied Environmental and Social Education Research, Queensland University of Technology.
[Online: https://eprints.qut.edu.au/53908/1/53908.pdf#page=19]

Nolas, S. (2011). Grounded theory approaches. In N. Frost (Ed.), *Qualitative research methods in psychology – Combining core approaches* (pp. 16–43). Open University Press.

Özmen, Z. M., & Baki, A. (2018). The views of undergraduate students about their undergraduate statistics course process. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS-10, July, 2018)*, Kyoto, Japan. International Statistical Institute.
[Online: https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_3E3.pdf?1531364258]

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307–332.
[Online: https://doi.org/10.3102/00346543062003307]

Pfannkuch, M., & Wild, C. J. (2000). Statistical thinking and statistical practice: Themes gleaned from professional statisticians. *Statistical Science*, *15*(2), 132–152.
[Online: https://projecteuclid.org/download/pdf_1/euclid.ss/1009212754]

Piaget, J. (1967). *Six psychological studies*. Vintage Books.

Picasso, P. (1935). *Head of a woman*. [Oil on Canvas]. Sotheby's, London. Licensed under Creative Commons 2.0. https://creativecommons.org/licenses/by/2.0/.
[Online: https://www.flickr.com/photos/gandalfsgallery/27485858810]

Primi, C., Donati, M. A., & Chiesi, F. (2016). A mediation model to explain the role of mathematics skills and probabilistic reasoning on statistics achievement. *Statistics Education Research Journal*, *15*(2), 169–178.
[Online: https://iase-web.org/documents/SERJ/SERJ15(2)_Primi.pdf]

Reid, A. & Petocz, P. (2002). Students' conceptions of statistics: A phenomenographic study. *Journal of Statistics Education*, *10*(2).
[Online: https://doi.org/10.1080/10691898.2002.11910662]

Rokeach, M. (1968). *Beliefs, attitudes, and values: A theory of organization and change*. Jossey-Bass.

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

Saldaña, J. (2015). *The coding manual for qualitative researchers*. SAGE Publications.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, *22*(2), 63–75.
[Online: https://doi.org/10.3233/EFI-2004-22201]

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* SAGE Publications.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.
[Online: http://www.jstor.org/stable/1738360]

van Rjin, Rembrandt. (1632-1636). *Rembrandt wearing a soft cap: Full face: Head only*. [Etching on paper]. Statens Museum for Kunst. Licensed under Creative Commons 1.0. https://creativecommons.org/publicdomain/zero/1.0/.
[Online:https://www.europeana.eu/portal/en/record/2020903/KKSgb8922.html?q=Rembrandt#dcId=1571637651476&p=7]

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223–248.
[Online: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1751-5823.1999.tb00442.x]

Wild, C. J., Utts, J. M., & Horton, N. J. (2018). What is statistics? In D. Ben-Zvi, K. Makar, & J. B. Garfield (Eds.), *International handbook of research in statistics education* (pp. 5–36). Springer International Publishing.

Zieffler, A., Garfield, J., delMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, *7*(2), 40–58.
[Online: https://iase-web.org/documents/SERJ/SERJ7(2)_Zieffler.pdf?1402525008]

Zieffler, A., Park, J., Garfield, J., delMas, R., & Bjornsdottir, A. (2012). The Statistics Teaching Inventory: A survey on statistics teachers' classroom practices and beliefs. *Journal of Statistics Education*, *20*(1), 1–29.
[Online: https://amstat.tandfonline.com/doi/abs/10.1080/10691898.2012.11889632]

NICOLA JUSTICE
Mathematics Department
Pacific Lutheran University
1010 South 122nd Street
Tacoma, WA 98447
United States

APPENDIX A: SURVEY INSTRUMENT

*Data from the last two items were not part of this analysis but are included to illustrate the context of the survey.

The following seven questions are designed to help convey your impressions of statisticians and statistics. Please communicate clearly, although complete sentences are not necessary (e.g., bulleted lists may be appropriate, etc.)

1. What are 2-4 typical tasks a statistician might do in their work? (open-ended short essay response)
2. List 2-4 skills or abilities a successful statistician needs to have. (open-ended short essay response)
3. List 2-4 character traits a successful statistician needs to have. (open-ended short essay response)
4. How would someone recognize if statistical methods are useful? (open-ended short essay response)
5. What distinguishes statistics from math? (open-ended short essay response)
6. What distinguishes statistics from other natural sciences (e.g., biology, chemistry, physics)? (open-ended short essay response)
7. If a stranger on the bus asked you, "what is statistics?" - What would you say? (open-ended short essay response)

The next two questions (and the final questions of this survey) are regarding statistical content. Please read each question carefully and choose the best response.

1. It has been established that under normal environmental conditions, adult largemouth bass in Silver Lake have an average length of 12.3 inches with a standard deviation of 3 inches. People who have been fishing Silver Lake for some time claim that this year they are catching smaller than usual largemouth bass. A research group from the Department of Natural Resources took a random sample of adult largemouth bass from Silver Lake. Which of the following provides the strongest evidence to support the claim that they are catching smaller than average length (12.3 inches) largemouth bass this year?

   a. A random sample of a sample size of 100 with a sample mean of 12.1.
   b. A random sample of a sample size of 36 with a sample mean of 11.5.
   c. A random sample of a sample size of 100 with a sample mean of 11.5.
   d. A random sample of a sample size of 36 with a sample mean of 12.1.
   e. More than one of the above are equally correct to me.

2. A game company created a little plastic dog that can be tossed in the air. It can land either with all four feet on the ground, lying on its back, lying on its right side, or lying on its left side. However, the company does not know the probability of each of these outcomes. Which of the following methods is most appropriate to estimate the probability of each outcome?

   a. Because there are four possible outcomes, assign a probability of 1/4 to each outcome.
   b. Toss the plastic dog many times and see what percent of the time each outcome occurs.
   c. Simulate the data using a model that has four equally likely outcomes.

APPENDIX B: LIST OF GROUNDED-THEORY THEMES PREVALENT IN 10% OR MORE OF PARTICIPANTS

Themes are given by order of prevalence of students who referenced them. Where there was a tie, they are presented alphabetically.

Theme	Prevalence	Description	Examples
Logic/Math	66%	References to needing math skills, mathematical reasoning, or logic skills. *Note: references to statistics being a subset of mathematics has its own category	“logical thought process” “(statistics is) a mathematical method for finding past trends or expected trends in data.”
Real-World Problem-Solving	66%	References application to the real world, solving real-world problems, or statistics being relevant to everyday life	“Uses of data and real-world application to business” “organization of data with real life” “...uses data to make conclusions about the real world.” “analyze financial performance, create econometric models to explain a variable and make predictions”
Analyze Data	59%	Uses "analyze data" or a derivation	“analyze the data,” “analysis”
Collect Data	57%	Uses "collect data" or a derivation	“gather data,” “data collection”
Patterns	48%	References to statistics involving finding or seeing patterns, correlations, trends, or relationships.	“A mathematical method for finding past trends or expected trends in data.” “find/make correlations”
Buzzwords	45%	Use of any of the following terms (precisely, with or without elaboration): analyze, *p*-value, *z*-score, significant (or significance).	“conduct an experiment, find p-scores, generate random distributions, calculate z-scores” “statistics is about the interpretation of data and its significance”
Tests and Procedures	43%	Includes performing tests (either generally or by specific name) or alludes to conducting (routine) analyses.	“calculate *p*-values and *z* scores...etc.” “Perform regressions, graphically model data, create confidence intervals” “...test to see if information is significant” “...the use of experiments, surveys, or other simple tests.”
Communication	41%	References to communication or presenting findings to others. *Note, does not include interpretation	“present data / findings” “format that information into understandable forms”
Technology	41%	References to the general use of technology or software, or to specific technology tools (e.g., R statistical software).	“Know how to use statkey [*sic*]” “computer skills” “Proficiency in multiple programming languages”
Variation and Uncertainty	39%	References to uncertainty, likelihood, chance, probability, or risk	“If you can use data to try to predict future outcomes or predict how likely previous outcomes were.” “In stats nothing is proven and there are a lot of outside factors that can affect our conclusions.” “Statistics helps determine whether the result of an experiment occurred by chance or not.”
Perseverance	34%	References to stretching outside of what is comfortable or natural: e.g., persistence, determination, patience	“determination” “able to focus for long periods of time” “patient and dedicated”

Subset	34%	Subset or branch of mathematics or of another specific type of another discipline or of sciences more generally	"A mathematical method for finding past trends or expected trends in data." "...statistics is a part of applied mathematics"
Conclude	30%	Uses "conclude" or a derivation of "conclude" or references to making a determination or decision. *Note, does not include interpret.	"drawing conclusions from data" "perform statistical analysis of some form on data, and/or determine the significance of their findings."
Critical/Analytical Thinking	30%	References to thinking critically or critiquing to make decisions. *Note, does not include analysis or analyze data.	"critical thinking" "Decide the best way to frame a survey's questions"
Detail-Oriented	30%	References to paying close attention, or noticing or observing things others might overlook	"attention to detail," "eye for detail" "critical examination," "observational"
Interpret	30%	References specific to interpreting data or results	"Statistics is about the interpretation of data and its significance." "Be able to read a graph, be able to interpret many different scores."
Ethics and Integrity	18%	References to avoiding personal bias, reporting with integrity, or using ethical practices	"non-biased attitude" "integrity"
People Skills	18%	Any reference to working with people, not including communication to people, etc., but rather more general people skills	"Works well with others..." "Teamwork..."
Generalize	16%	Generalizing to a population or process	"Collecting and analyzing large numbers of data and applying it to a larger population or bigger underlying method" "Ability to take sample data and determine whether it can be generalized to a larger population"
Organized	16%	Reference to being organized as a general skill or character trait. *Note, does not include references to organizing data, etc.	"organization skills" "organized"