The Impact of Using Relevant Context on Student Comprehension and Attitude in a Collegiate Introductory Statistics Unit on Probability

Ryan F. Rzeszutko  
*University of Missouri-St. Louis, ryan.rzeszutko@umsl.edu*

Jennifer L. Petrie  
*University of Missouri-St. Louis, shrenskerj@umsl.edu*

Dwayne T. James  
*University of Missouri-St. Louis, jamesdt@umsystem.edu*

Follow this and additional works at: [https://irl.umsl.edu/dissertation](https://irl.umsl.edu/dissertation)

Part of the Science and Mathematics Education Commons

**Recommended Citation**

[https://irl.umsl.edu/dissertation/1307](https://irl.umsl.edu/dissertation/1307)

This Dissertation is brought to you for free and open access by the UMSL Graduate Works at IRL @ UMSL. It has been accepted for inclusion in Dissertations by an authorized administrator of IRL @ UMSL. For more information, please contact marvinh@umsl.edu.
The Impact of Using Relevant Context on Student Comprehension and Attitude in a Collegiate Introductory Statistics Unit on Probability

Dwayne T. James, P.E.
M.S. Civil Engineering, University of Missouri-Rolla, 1997
B.S. Civil Engineering, University of Kansas, 1994

Jennifer Petrie
M.A. Mathematics, Washington University in St. Louis, 2004
B.S. Mathematics, University of Missouri–Columbia, 2000
B.E.S. Curriculum and Instruction, University of Missouri–Columbia, 2000

Ryan Rzeszutko
M.S. Mathematics, Southern Illinois University Edwardsville, 2014
B.S. Mathematics, University of Illinois at Urbana-Champaign, 2010

A Co-Authored Dissertation submitted to
The Graduate School at the University of Missouri-St. Louis
in partial fulfillment of the requirements for the degree
Doctor of Education with an emphasis in Educational Practice

May 2023

Dissertation Committee

Dr. Keith W. Miller, Ph.D.
Chairperson

Dr. Charles R. Granger, Ph.D.

Dr. Helene J. Sherman, Ed.D.

Copyright, Dwayne T. James, Jennifer Petrie, and Ryan Rzeszutko, 2023
ABSTRACT

The typical collegiate introductory statistics course poses significant challenges for students. Many do not fully comprehend key course skills, and it is common for students to exit the class with a neutral or negative attitude toward statistics. To measure the impact of using relevant contextual examples as an instructional strategy during a probability unit, in-class activities were designed to align with areas of interest for participants as identified by a student interest inventory. It was hypothesized that the use of relevant context would create a significant difference in the comprehension or attitude of students enrolled in an introductory statistics course at an urban midwestern university. Following a quantitative analysis of comprehension and attitude, interviews and focus groups were conducted with students from both the treatment and control groups to better understand the factors that influence student comprehension and attitude. Quantitative results reveal no significant differences in either comprehension or attitude as a result of the relevant context. Qualitative findings suggest other factors such as the instructor, class structure, and previous experiences play a larger role in shaping student comprehension and attitude.

Keywords: relevant context, comprehension, attitude, statistics, probability, statistics education, college instruction, Situated Learning Theory, Realistic Mathematics Education
ACKNOWLEDGEMENTS

In the words of Dr. Granger, “this section proves” our gratitude for the support of our dissertation committee. The guidance of Dr. Keith Miller, Dr. Charles Granger, and Dr. Helene Sherman helped turn our three individual burning questions into a cohesive, full-fledged dissertation. We appreciate your time and effort invested into us.

To my partner Freddie, who provided continued encouragement. To my parents Edward and Sharon and brother Darren, who have provided guidance in my every endeavor. To my grandparents, your love continues to be my inspiration. To my family and friends, thank you for supporting me throughout this journey. D.T.J.

To my daughters Whitney and Annalie, who I love more than anything. I want the two of you to know that you can achieve anything you set your mind to doing and that it’s never too late in life to chase a dream. To my parents Vicki and Steve, for their unwavering support through all of life’s challenges. J.P.

To my mother Janet, who convinced me to start my doctorate. To my wife Jordanna, who supported me throughout. To my daughter Ellis, who motivated me to finish. And finally, to my father Ralph, who would have been quite proud. R.R.
Table of Contents

CHAPTER

1 INTRODUCTION ________________________________________________7

Background____________________________________________________7

Problem Statement_____________________________________________9

Purpose_______________________________________________________11

Research Questions____________________________________________11

Hypotheses____________________________________________________12

Definitions of Terms____________________________________________13

Procedures____________________________________________________14

Significance____________________________________________________15

Summary_______________________________________________________16

2 LITERATURE REVIEW _________________________________________17

Introduction____________________________________________________17

Literature Search Strategy_______________________________________18

Theoretical Framework__________________________________________18

Review of Literature____________________________________________22

Summary_______________________________________________________42

3 METHODOLOGY ______________________________________________43

Appropriateness of the Research Design____________________________43

Research Questions and Hypotheses_______________________________44

Research Design________________________________________________45
APPENDIX ITEMS

A LOGISTICS

A.1 - IRB Consent Form

A.2 - Volunteer / Extra Credit Opportunity Form

B RELEVANCE INSTRUMENTATION AND DATA

B.1 - Student Interest Inventory

B.2 - Interest Inventory Results and Narrative

B.3 - Exit Slip Form and Exit Slip Data

C INSTRUMENTS

C.1 - Comprehension Pretest

C.2 - Posttest and Validation Data

C.3 - SATS-36 Explanation

C.4 - Qualitative Script / Interview Questions

D INSTRUCTIONAL MATERIALS

D.1 – Treatment Group Activities

D.2 – Control Group Activities

E ADDITIONAL DATA

E.1 - Demographic Tables

E.2 - Attendance Records

E.3 - Final Code Book

F INSTRUCTOR REFLECTIONS

F.1 - Selected Researcher Comments

F.2 - Selected Researcher Journal Entries
Chapter 1

Introduction

Background

When compared to other branches of mathematics education, the field of statistics instruction is not as well researched or as established (Garfield & Ben-Zvi, 2007). Advances in technology that began during the 1990’s enabled instructors to shift from a primarily mechanical and computational course to one that was focused on holistic understanding and decision making (Pfannkuch & Wild, 2000). The study of student attitude toward statistics rose into prominence during the 1990’s, largely thanks to the work of Gal et al. (1997). The Guidelines for Assessment and Instruction in Statistics Education project (GAISE) that was completed in 2005 emphasized the need for greater conceptual understanding and collaboration in statistics courses (GAISE, 2005).

Despite the ambitious and wide-reaching changes proposed by the GAISE committee, enhanced student learning outcomes have been difficult to achieve on a consistent basis. Numerous studies point to the difficulty that students have remembering and applying statistical skills learned during an introductory statistics course (Garfield & Ben-Zvi, 2007; Konold, 1995). Many introductory statistics students do not exhibit growth in their ability to think and reason using statistics and probability (delMas et al., 2007; Trumpower, 2010). The attitude of students toward statistics frequently remains unchanged, or even worse, deteriorates over the course of an introductory statistics class (Schau & Emmioglu, 2012). While considerable progress has been made rethinking the purpose of statistics education, at present there is a definite need for instructional
strategies that can provide lasting improvement to student learning outcomes (Garfield & Ben-Zvi, 2007).

Within the introductory statistics course, the concept of probability presents an imposing barrier to success. Student intuition regarding chance outcomes is frequently incorrect, leading to conflicting messaging and confusion in learning required skills (Garfield & Ben-Zvi, 2007). Students frequently struggle to apply concepts outside of the exact context in which they were learned (Chow & VanHaneghan, 2016), and sometimes completely fail to recognize that probability is required to answer a problem (Uegatani et al., 2020). A deficiency in probabilistic reasoning is especially troublesome in introductory statistics because the vast majority of statistical inference is framed in terms of $p$-values and probability. Academic challenges with probability follow students throughout the remainder of the course, inevitably leading to further frustration and gaps in understanding (Cobb, 2015).

To address this continued shortcoming in statistics education, many have offered their guidance regarding the reconstruction of the statistics curriculum. The GAISE Committee reconvened in 2016 and found that their initial recommendations were still relevant. One initial recommendation by the committee, “to use real data” (GAISE, 2005, p. 16), was reworded for additional clarity and precision to “integrate real data with a context and a purpose” (GAISE, 2016, p. 3). Statistics education expert, George W. Cobb, recommended that instructors “exploit context” to enhance student interest and motivation, and that teachers should “teach through research” (Cobb, 2015, pp. 276-277) to provide students with the necessary experience to analyze authentic data of their choosing to master course objectives.
Many other pedagogical recommendations by GAISE and Cobb have been researched at length. In comparison to traditional lecture-based instruction, other instructional strategies such as active learning have been studied extensively and consistently find positive results on student comprehension (Freeman et al., 2014). Surprisingly, research also suggests that the use of active learning is not always sufficient to change student attitudes toward statistics, even when compared to traditional teaching methods (Bateiha et al., 2020). It can be difficult to change the attitude of a student toward a discipline (Schau & Emmioglu, 2012), but finding a way to do so is critical for improving statistical competency.

Several important gaps in statistics education need to be addressed in order to improve student comprehension and promote positive shifts in attitude toward the discipline. Both George W. Cobb and the GAISE 2016 committee recommended the use of relevant context in statistics, because its use has the potential to improve student learning outcomes (Cobb, 2015; GAISE, 2016). Few studies have attempted to measure the impact of relevant context alone, and of those that do, many compare the context in an active learning setting to a traditional lecture-based control. The goal of comparing an active learning classroom setting using relevant context within situated experiences to an active learning classroom that utilizes generic context is to come closer to isolating the impact of relevant context related to the two desired student learning outcomes of comprehension and attitude.

**Problem Statement**

As the need for statistical literacy has increased and permeated numerous occupations over recent years, the number of college students who have enrolled in an
introductory statistics course as part of their degree requirements has also risen (Onwuegbuzie & Wilson, 2003). For many students, an introductory statistics course serves as their only exposure to formal statistical reasoning (Ramirez et al., 2012). As a result, the statistical competency of our college student population is heavily reliant upon the learning outcomes achieved during this course (Ramirez et al., 2012). Nationwide testing has revealed that college students who enrolled in introductory statistics courses did not improve in key areas such as probabilistic reasoning and statistical inference upon completion of their statistics course (delMas et al., 2007). Additionally, many students considered their introductory course to have been a negative experience (Onwuegbuzie & Wilson, 2003). Attitude was found to be a significant predictor of comprehension, specifically statistical literacy (Hay et al., 2015). Because of this connection between attitude and statistical literacy, it is important for researchers to not just focus on one of these learning outcomes. Instead, attitude and statistical literacy should be considered jointly.

Currently, little research focuses specifically on collegiate content delivery methods that are employed in statistics courses (Bateiha et al., 2020). The use of relevant context and real-life data deviates from the typical structure of a collegiate statistics course, yet has the potential to spark student interest and motivation in the content area (Cobb, 2015). Evidence also suggests that the use of relevant context fosters a deeper level of statistical thinking, which could lead to higher levels of comprehension and retention (Pfannkuch, 2011). More research is needed to isolate the impact of relevant context on student achievement and attitude, particularly at the collegiate level.
Purpose

The content delivery was modified through the customization of class problems to accommodate student interests, as related to two key student learning outcomes to determine the impact of relevant context on the understanding of probability concepts in a unit from a college introductory statistics course. First, student comprehension of unit skills was assessed on pre- and post-unit exams. Second, a pre- and post-version of the Survey of Attitudes Toward Statistics (SATS-36) instrument developed by Candace Schau (2003) measured student change in attitude toward statistics. Additionally, one on one interviews and focus groups were conducted with selected students to find out in what ways students describe the factors that affected their comprehension of probability topics and their attitude toward statistics based on the probability unit.

Research Questions

To test the effects of using relevant context during the probability unit in an introductory statistics course, the following three questions were raised:

1. To what extent does using relevant context within situated experiences affect student comprehension of probability concepts?

2. To what extent does using relevant context within situated experiences affect student attitude toward statistics?

3. In what ways do students describe the factors that affected their comprehension of probability and their attitude toward statistics based on the probability unit?
Hypotheses

Student learning outcomes were compared based upon two different types of context delivered within the probability unit of an introductory statistics course.

- Students in the treatment group engaged in active learning activities centered around relevant context related to student interests, as identified through a preliminary interest survey and verified through exit slip surveys.
- Students in the control group engaged in active learning activities centered around generic context that had not been tailored to student interests.

RQ#1  $H_01$: There is no significant difference in average comprehension gains between students in the treatment group and students in the control group.
$H_{A1}$: There is a significant difference in comprehension gains between students in the treatment group and students in the control group.

RQ#2  $H_02$: There is no significant difference in any of the six attitude components of Affect, Cognitive Competence, Value, Difficulty, Interest, or Effort between students in the treatment group and students in the control group.
$H_{A2}$: There is a significant difference in at least one of the six attitude components of Affect, Cognitive Competence, Value, Difficulty, Interest, or Effort between students in the treatment group and students in the control group.
Definitions of Terms

● *Attitude* toward statistics is a “multidimensional concept referring to distinct, but related dispositions pertaining to favorable or unfavorable responses with regard to statistics and statistics learning” (Vanhoof et al., 2011, p. 35)

● *Comprehension* refers to how well students understand educational concepts. Comprehension for an introductory statistics course includes the aspects of statistical literacy, statistical reasoning, and statistical thinking (Garfield et al., 2006). Only two components of statistical literacy and statistical reasoning were considered because statistical thinking cannot be assessed effectively without the use of open-ended questions.

● *Context* is what gives a mathematics or statistics problem meaning or describes the general setting in which the problem occurs.

● *Learning Outcomes* are defined as the changes in comprehension of probability unit skills and changes in attitude toward statistics for individual students in the course.

● *Relevant Context* refers to a context which “relates to prior experience; elicits positive emotional reactions; or is associated with feelings of worth, value, or is connected to important personal goals” (Walkington & Bernacki, 2018, p. 50).

● *Situated Experiences* are the in-class, interactive activities developed for the intervention. These activities were grounded in relevant or non-relevant contexts and were based on a combination of Jean Lave’s Situated Learning Theory and Hans Freudenthal’s Realistic Mathematics Education.
• **Statistical Literacy** includes understanding terminology, notation, and various visual representations of data, along with basic mathematical computations that stem directly from a specific formula. Garfield et al. (2006) states that it is “understanding and using the basic language and tools of statistics” (np).

• **Statistical Reasoning** includes relating statistical ideas, using statistical concepts to solve more complicated problems, inferring meaning from computations, and discriminating between different types of statistical problems. Garfield and Chance (2000) state that statistical reasoning is “the way people reason with statistical ideas and make sense of statistical information” (np).

**Procedures**

Using a quasi-experimental, mixed methods approach, insight was gained on the impact that utilizing relevant context as a basis for an instructional strategy has on student comprehension of probability and student attitude toward statistics. Student comprehension was measured quantitatively through the use of a pretest and posttest to be administered at the beginning and end of a probability unit. Change in attitude was also measured quantitatively through a pre- and post-version of the *SATS-36*. Additionally, a thematic understanding of the ways students describe their change in comprehension and attitude was examined through a qualitative lens. Through the use of targeted focus groups and individual interviews, student perceptions were explored regarding their experience with relevant context.
Significance

Probability concepts have consistently been an area of difficulty for students to master (delMas et al., 2007; Garfield, 2002). Although researchers have attempted to mitigate this problem using different teaching strategies and techniques, there have been varying results of effectiveness. In 2016, the GAISE committee recommended that instructors “integrate real data with a context and a purpose” (p. 3), and the results from this study can be used to assess the validity of this recommendation. Additionally, the knowledge gained by determining the influence of teaching with relevant contextual examples on student comprehension and attitude is of benefit to all instructors teaching probability units.

On a local level, one of the authors of this dissertation has been coordinating an introductory course in probability and statistics for the past 15 years. Records from Spring 2011 through Spring 2020 (ignoring examination results during the COVID-19 pandemic due to pandemic-related inconsistencies) provide a decade of historical data for face-to-face students enrolled in this course and confirms that this student population has also struggled with topics in probability. Probability concepts were tested throughout all of Exam 2 and on half of Exam 3 during this 10-year period (see Table 1.1).

Table 1.1

| 10-year Historical Exam Average in the Introductory Statistics Course |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
|                            | Exam 1  | Exam 2  | Exam 3  | Exam 4  |
| Average Score              | 86.3%   | 70.6%   | 72.0%   | 77.7%   |
| (n=1094)                    | (n=1059)| (n=895) | (n=929) |

Note. Data is from students who were enrolled in sections taught by one of the researchers.
The average for those tests which cover probability topics are considerably lower than those which cover other statistical concepts. For this reason, it is important to discover ways to help students learn probability concepts as successfully as they have mastered other statistical concepts. In its historical format, course instructors use traditional textbook contexts such as coins, dice, decks of cards, and balls in urns to explain many probability concepts. Instructors participating in the intervention used classroom examples involving relevant contextual data. While there is no historical data on attitude for this student population, the present study considered the effect that context has on these measures. The overall results will be used to inform teaching practices for the probability unit in this course.

Summary

Students have historically struggled in their collegiate introductory course in probability and statistics. Research in this area is relatively new compared to other fields of mathematics, and not much is known about the impact of incorporating relevant context on student comprehension and attitude. This quasi-experimental, mixed methods study was intended to increase the general knowledge of how relevant context influences comprehension and attitude in a collegiate introductory statistics course, to quantify the impact of relevant context on student learning outcomes and gain insight into the primary influences on student comprehension and attitude throughout this probability unit.
Chapter 2

Literature Review

Introduction

For many college students, an introductory statistics course presents significant challenges (delMas et al., 2007; Garfield, 1995; Konold, 1995). Students often struggle to comprehend key course concepts (delMas et al., 2007; Garfield, 2002), which prevents them from being able to apply those skills in the future. In addition, students frequently develop a negative attitude toward statistics (Schau & Emmioglu, 2012). A negative attitude, particularly regarding personal interest toward statistics, is associated with a more pessimistic view of personal ability to succeed in the course (Carmichael et al., 2010; Hay et al., 2015). Students with a negative attitude toward a course are also less likely to apply skills from the course later on in life (Bandura, 1997; Di Martino & Zan, 2010). Because the majority of students do not take more than one statistics course, the statistical competency of our society is dependent on the learning outcomes achieved at the introductory statistics level (Ramirez et al., 2012).

Over the past 30 years, a statistics reform movement has taken place at the collegiate level to address some of these issues. Research resulting from these reforms has shown that alternative teaching strategies, such as situating learning within the context where it is applied in the real world, enhance student performance in these classes (Garfield & Ben-Zvi, 2009; Moore, 1997). Another strategy includes working with contextual, real-life data. This use of context has been shown to increase student interest and achievement, as it draws on student background knowledge from outside the
Presently, there is still limited research on instructional strategies in collegiate introductory statistics courses (Bateiha et al., 2020). Given the value of student comprehension and attitude toward statistics, more research is needed on how instructors can improve these learning outcomes for students. Part of this gap in literature was addressed by attempting to isolate the impact of relevant context on student comprehension and attitude.

**Literature Search Strategy**

The search strategy was established using a literature review outline based on keywords. Keywords included, but were not limited to realistic mathematics, relevant context, situated learning, statistics, probability, mathematics, cognitive domain, content mastery, comprehension, achievement, affective domain, and attitude. Google Scholar and the University of Missouri Library were leveraged for these searches. As part of the literature review, 78 articles, dissertations, books, and reports were reviewed and sorted into five main themes: Situated Learning Theory, Realistic Mathematics Education, statistical context, statistical comprehension, and statistical attitude.

**Theoretical Framework**

To promote student comprehension of statistical content, as well as student attitude toward statistics, the use of relevant context was examined as an instructional strategy. The incorporation of relevant context into a probability unit was guided by two educational theories. Situated Learning Theory (SLT), created by Jean Lave and Etienne Wenger, and Realistic Mathematics Education (RME), created by Hans Freudenthal,
were merged to form the framework for facilitating student learning outcomes within the setting of a collegiate introductory statistics course.

**Situated Learning Theory**

Situated Learning is the experiential learning grounded in real-life context that takes place in the environment in which it is applied. Jean Lave states "that learning as it normally occurs is a function of the activity, context, and culture in which it occurs" (Kearsley, 2011, np). According to SLT, learning is more effective when individuals actively participate in the process of knowledge acquisition (Lave & Wenger, 1991). Rather than passively learning through listening to someone else, learners should take a hands-on role and should actively be challenged to make decisions and justify their reasoning. Additionally, according to Lave & Wenger (1991), learning can only take place when situated in the environment in which it is applied. For example, a carpenter will only truly learn how to create a chair through the act of creating one. Within a traditional classroom, the learning process can still be inspired by SLT, despite teaching skills outside of the environment where they will ultimately be applied. Material should be presented in rich context depicting where it will be used, and students should be given freedom to explore their own ideas in a community of practice. Settings that are too restrictive and do not allow students to engage authentically and learn from each other are not likely to lead to learning (Lave, 1991).

**Realistic Mathematics Education**

RME uses real-life experiences and the personal experiences of the student for developing mathematical concepts. In this theory, relevant situations that are realistic and relatable are highlighted in the learning process. RME “gives great attention to the
cognitive development of students” (Darto, 2021, p. 6). Students have a greater opportunity to learn when they are able to connect to the material and understand its importance to their life. The RME approach provides a more relatable environment for the students to learn while giving them an active role in the instructional process (Zakarita & Syamaun, 2017).

**Relevant Context**

In a SLT and RME learning environment, students can improve their own comprehension and attitude. Linking instructional materials to life experiences provides an opportunity for students to use their current knowledge and see themselves in the problems they will solve (Darto, 2021). Individuals learning with relevant context can better interpret the problem which enables them to relate to and actively engage with the material, as the problem more closely identifies with the environment and prior knowledge experience of the individual (Ulandari et al., 2019).

Both SLT and RME are united in their emphasis on the importance of relevance for student learning. Lave and Wenger (1991) argue that “there is no activity that is not situated” (p. 33) and that learning can only take place within the context of the actual environment where it will be applied. Freudenthal’s RME emphasizes the importance of grounding mathematics in reality and its uses in the real world (Van den Heuvel-Panhuizen, 2020). The use of relevant context fits the goals of both of these theories, provided that it is done correctly.

Care must be taken so that the context is either *authentic* (actual data from real-life) or *realistic* (fictional data but plausible). Many times, context serves as mere decoration for a mathematical problem, which does little to enhance the student
experience (Vos, 2020). By striving to create context that is either authentic or realistic, the instructor is providing students with a setting that can motivate learning and facilitate understanding.

**Figure 2.1**

*Visual Display of the Theoretical Framework*

As shown in Figure 2.1, the premises of SLT and RME inspired the use of relevant context to enhance student learning outcomes. According to these theories, active learning and relevant context can have a positive impact on student attitude and comprehension. Participation in active learning experiences, a core component of both SLT and RME, has been linked to numerous positive student learning outcomes. The use of active learning can lead to increased comprehension and retention, enhanced development of critical thinking skills, and improved attitude toward the content (Garfield & Ben-Zvi, 2009). The incorporation of relevant context can improve student motivation for content mastery and provide an outlet for increased communication regarding mathematical findings (Darto, 2021). When problems are centered on realistic situations and involve hands-on experience, students show both higher levels of
comprehension and improvement in their attitude toward the course (Pinwanna, 2015). Context tied to social issues can also serve as a source of engagement for students, serving as a springboard for future development of inferential reasoning skills (Zapata-Cardona, 2018).

**Review of Literature**

*A History of Statistics Reform*

College students have historically struggled with introductory probability and statistics courses (Chow & Van Haneghan, 2016; delMas et al., 2007; Garfield, 1995, 2002; Garfield & Ben Zvi, 2007; Konold, 1995; Pollatsek et al., 1987). Students frequently fail to comprehend key concepts within the course and come away with a negative attitude toward the discipline of statistics (Schau & Emmioglu, 2012). To address these struggles, a reform movement began in the early 1990’s.

During June of 1990, the Workshop on Statistical Education was held at the University of Iowa. Thirty-nine educators and professionals came together to identify the issues surrounding statistics education and to formulate suggestions to address these challenges (Cobb, 1992). The result of this conference was a report detailing 17 topics that were important to teach in an introductory statistics course listed by level of importance, nine suggestions that professors could use to aid in their teaching and students could use to aid in their learning, along with numerous other suggestions for administration, institutions, and professional organizations (Cobb, 1992). The eighth and tenth of these curricular suggestions were about basic distributions such as the normal and binominal distribution and elementary probability, including tree diagrams and conditional probabilities, respectively. The third of these pedagogical recommendations
for teachers stated, “Students rely too much on formulas rather than on thinking. Do everything possible to improve a student's self-esteem and his or her appreciation of and self-confidence in statistical studies” (Cobb, 1992, p. 41).

In 1991, the board of governors of the Mathematics Association of America issued a “call for change” in mathematics education (Cobb, 1992). One response to this call was a focus group moderated by George W. Cobb that dealt with many of the topics brought up at the 1990 workshop. Out of this focus group came three broad recommendations which could be applied to any introductory statistics course. The first of these recommendations was that statistics courses should emphasize statistical thinking by using data, understanding the importance of data production and study design, understanding that variability is always present, and that variability can be quantified and explained (Cobb, 1992). The second recommendation was to teach with more data and fewer concepts and theories, because as Cobb explained, “statistical concepts are best learned in the context of real data sets” (Cobb, 1992, p. 7). Part of this recommendation was to rely heavily on real-life data, not just realistic data, to rely on computers rather than formulas, and to place more of the emphasis on the statistical concepts rather than computational fluency (Moore, 1997). The final recommendation was to foster active learning in any of the following ways: by using small groups for discussion of ideas and to solve problems, by having a laboratory component to the course so that students participate in data generation, by using demonstrations with class-generated data when individual or small group hands-on activities cannot effectively be done, by assigning written or oral presentations so that students have an opportunity to explain their thinking, and by assigning either small group or individual projects (Cobb, 1992).
For the next decade, the ideas that came out of this workshop and focus group began to lay the foundation for changes at the collegiate level. Changes were made to content, pedagogy, and technology (Moore, 1997). There became a wider consensus about what topics ought to be taught in an introductory statistics course and what topics should be de-emphasized (Scheaffer, 2001). The proposals from George W. Cobb also formed the basis of the Guidelines for Assessment and Instruction in Statistics Education (GAISE) Project in 2005 (GAISE College Report, 2005; Tu & Snyder, 2017), which outlined six recommendations for introductory statistics courses. These six suggestions had a profound impact on the statistics reform movement (GAISE College Report ASA Revision Committee, 2016), and were re-evaluated in 2016. The original six recommendations largely withstood the test of time during the evaluation and were only refined slightly. After the modifications in 2016, the six GAISE recommendations were: “Teach statistical thinking; Focus on conceptual understanding; Integrate real data with a context and purpose; Foster active learning; Use technology to explore concepts and analyze data; and Use assessments to improve and evaluate student learning” (GAISE, 2016, p.3). Three of the suggestions from this list, integrating real data with a context and purpose, fostering active learning, and using assessments to evaluate student learning are ongoing areas of study and fit within the objectives of this dissertation.

**The Use of Relevant Context**

A context is relevant “if it relates to prior experience; or if it elicits positive emotional reactions; or if it is associated with feelings of worth, value, or is connected to important personal goals” (Walkington & Bernacki, 2018, p. 7). Since the initial 2005 GAISE report, researchers have called for using relevant contexts to teach statistical
concepts (Gould, 2010; Walkington & Bernacki, 2018). Limited research has been conducted on the impact of relevant context on student learning outcomes (Pfannkuch, 2011). Of the existing research, some studies show evidence of student comprehension gains when exposed to relevant context (Walkington & Bernacki, 2018), while others such as Hogheim & Reber (2015, 2017) reveal no significant improvement in student performance. These contradictory findings are due in part to a lack of consensus on what constitutes relevant context (Walkington & Bernacki, 2018). Before an effective intervention can be developed, it is essential to explore the concept of relevant context in more depth.

Numbers used for analysis in statistics must contain background and context to provide meaning for students - it is at this point that “numbers” turn into “data” (Cobb, 1992). Data can be separated into two distinct categories within the context of probability and statistics (Neuman et al., 2013; Singer & Willett, 1990). First, artificial data can be defined as hypothetical, simulated data designed for computational efficiency. Although the data may appear “real” at a glance, the contrived nature of the problem can become apparent upon investigation, frequently leading to a decrease in student attitude and motivation toward statistics (Singer & Willett, 1990). On the other hand, real-life data is authentic and is collected from classroom activities, research projects, or outside archives. Real-life data is more meaningful and intrinsically interesting to students, and can help spark student motivation to learn statistical concepts (Chick & Pierce, 2010; Neuman et al., 2013).

According to Walkington and Bernacki (2018), students enter the classroom with a “fund of knowledge” (p. 52) constructed from personal experiences. This fund of
knowledge already consists of experiences with data generated from sources such as social media, smartphone apps, pop culture, music, sports, video games, and much more (Gould, 2010; Walkington & Bernacki, 2018). Integrating these sources of context into classroom activities can ground the conceptual learning to an idea that students can easily relate to and understand (Walkington & Bernacki, 2018). In pulling from the class fund of knowledge, it is important to consider the level of contextualization. A shallow contextualization refers to actions such as swapping a generic problem context with one that is interesting to students, with little change to the initial structure of a problem. For example, instead of a word problem where a child is buying apples, they may be buying action figures from a popular video game instead. Such superficial changes can lead to an initial flash of interest in students but contribute little to student comprehension and may even serve as an unnecessary distraction (Vos, 2020; Walkington & Bernacki, 2018). In contrast, a deep contextualization places the concept within a situation for how it might be used in the daily life of a student. Students must develop a solution that has real-world value and authenticity (Vos, 2020; Walkington & Bernacki, 2018). If a student connects with the context of the problem, this deep contextualization is associated with higher buy-in and engagement for students (Clinton & Walkington, 2019).

In addition to the depth of the contextualization, it is also important to consider grain size and the amount of personalization of relevant context. Grain size refers to the broadness of appeal that a topic has for an audience. A topic with large grain size would have relevance to many students, while a small grain size would be relevant to only a few students (Walkington & Bernacki, 2018). Topics with larger grain sizes could appeal to larger audiences, but also could lack the degree of specificity that a smaller grain size
could provide. Whether grain size is large or small, there does not appear to be strong evidence of an advantage in long term student performance, as long as the context is personally relevant to the students (Clinton & Walkington, 2019; Walkington & Bernacki, 2018). Relevant context can be structured so that it is personalized for individual students. Students who were able to choose their context showed an increased performance on assessments than those who were randomly assigned a context (Clinton & Walkington, 2019). Clinton and Walkington (2019) caution against extensive personalization, as a student may lose interest in pursuing topics that are less relevant to their interests.

The Impact of Relevant Context on Student Learning Outcomes

In accordance with the theories of both RME and SLT, students should experience greater learning outcomes in settings that are authentic and relevant to their own interests. Existing research is far from conclusive on the exact impact of relevant context on student comprehension and attitude, particularly toward probability and statistics (Walkington & Bernacki, 2018). Many studies have uncovered evidence that the use of relevant context and real-life data leads to higher levels of comprehension of course objectives (Hay et al., 2015; Neuman et al., 2013; Papaieronymou, 2017). Others have linked the use of relevant context to higher levels of mathematical and statistical reasoning skills (Garfield & Ben-Zvi, 2009; Hogheim & Reber, 2017; Pfannkuch, 2011). Still others suggest that increased relevance leads to improved student attitude, particularly in their interest toward the subject (Neuman et al., 2013; Wathen & Rhew, 2019). Finally, other studies have suggested that relevant context can improve student recall (Singer & Willett, 1990), increase class participation and enjoyment (Naidu &
Sanford, 2017), and improve awareness over how course content is used in daily life (Daniel & Braasch, 2013).

Despite these promising findings, many other studies expose potential pitfalls when attempting to implement relevant context into activities. Real-life data can be messy by nature, and difficult to incorporate into assessments (Singer & Willett, 1990). The simple act of providing context does not give a task meaning in of itself. As Vos (2020) explains, there is no inherent need to know how many quarter pizzas fit inside three and a half pizzas. When the question posed has no inherent purpose other than to get the right answer, the context is simply decorative, and does not contribute to deeper learning (Vos, 2020). In such a scenario, it is unlikely that students would experience any improvement to learning outcomes. Uses of context, particularly when incorporating shallow contextualizations, can serve as a distraction for student learning (Clinton & Walkington, 2019). Interestingly, an activity that provokes too much student interest can also inhibit comprehension, as students get distracted from the skills intended to be developed in the activity (Clinton & Walkington, 2019; Nisbet et al., 2007). Finally, instructional design must be deliberate and thoughtful in ensuring that students can develop the desired mathematical and statistical skills. In a relevant and timely study on conditional probability and COVID-19 infection rates, high school students were engaged in an activity to determine whether it was just to require negative COVID tests for employment purposes. When presented with an open-ended challenge to justify their stance, many students opted not to use their recently gained knowledge in probability (Uegatani et al., 2020).
**Statistical Comprehension**

Measuring student comprehension in an introductory statistics course has shifted focus in meaning since the educational reforms took place. Prior to the reforms, most college statistics courses were merely asking students to manipulate formulas and perform calculations (Brown, 2019). However, the 2016 GAISE recommendation to teach statistical thinking has altered the focus of these collegiate introductory courses, requiring more of students than just computational fluency. Garfield et al. (2006) identified three aspects of student comprehension for an introductory course - statistical literacy, statistical reasoning, and statistical thinking. DelMas (2002) listed the various types of tasks that one might expect students to complete for each of these comprehension components (see Table 2.1). Instructors may ask questions in any of these three categories, and they may or may not include computations.

**Table 2.1**

*Assessment Tasks that may Distinguish the Three Instructional Domains (delMAS, 2002, p.6)*

<table>
<thead>
<tr>
<th>BASIC LITERACY</th>
<th>REASONING</th>
<th>THINKING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify</td>
<td>Why?</td>
<td>Apply</td>
</tr>
<tr>
<td>Describe</td>
<td>How?</td>
<td>Critique</td>
</tr>
<tr>
<td>Rephrase</td>
<td>Explain</td>
<td>Evaluate</td>
</tr>
<tr>
<td>Translate</td>
<td>(The Process)</td>
<td>Generalize</td>
</tr>
<tr>
<td>Interpret</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The first aspect of comprehension, *statistical literacy*, refers to the basic understanding and use of statistical terms, notations, and data representations (Garfield & Ben-Zvi, 2007). Students demonstrating statistical literacy can digest the material well
enough to later think about it and make informed decisions based on their understanding of information (Rumsey, 2002). Gould (2010) believed that statistical literacy is meant for the consumers of statistics, those who simply have a need to read and understand data. When considering content from a probability unit, statistically literate students should be able to read and use probability tables, perform basic probability calculations, understand the terminology used to express probabilistic ideas, and understand the meaning of their calculated probabilities. Statistical literacy typically serves as a foundation for statistical reasoning and statistical thinking, although some researchers have argued that reasoning and thinking can operate independently from literacy (Garfield, 2002; Garfield et al., 2006).

*Statistical reasoning* is the way in which students make sense of statistical ideas and information. It includes how they connect concepts together, explain processes, and interpret results (Garfield & Ben-Zvi, 2007). Statistical reasoning as it relates to probability may include “correctly using the ideas of randomness, chance, and likelihood to make judgements about uncertain events, knowing why not all outcomes are equally likely, and knowing when and why the likelihood of different events may be determined using different methods” (Garfield, 2002, np).

According to Garfield and Ben-Zvi (2007), there are many aspects to *statistical thinking*, including knowing how and why to use a particular technique, having a deep understanding of the theories behind various statistical processes, and understanding limitations and constraints. It also involves knowing how, when, and why simulations and inferential tools can be
used, and it involves using the context of a problem to plan investigations and draw conclusions. Gould (2010) believed that statistical thinking was meant for producers of statistics, those who would be performing experiments, interpreting, and analyzing their results, and sharing their conclusions to a larger audience.

**Challenges in Comprehension**

There are a number of misconceptions that students face when it comes to their comprehension concerning probability. Especially in an introductory statistics course, students struggle because the concepts taught often conflict with their intuition or previously held beliefs about a topic (Garfield, 1995). These intuitions are difficult to alter, even after presented with evidence (Konold, 1995). Furthermore, students can also hold many contradictory beliefs about a particular situation and apply their intuition in different ways when similar questions addressing identical content are phrased differently (Konold, 1995).

It has been well-documented that students have struggled with probability concepts for decades (delMas et al., 2007; Garfield & Ben-Zvi, 2007). Students tend to display incorrect reasoning skills about chance events (Garfield, 2003; delMas et al., 2007; Konold, 1995), randomness (Konold, 1995), the law of large numbers (Konold, 1995), and conditional probabilities and their applications (Chow & VanHaneghan, 2016; Konold, 1995; Pollatsek et al., 1987). Konold et al. (1993) noticed that how the question was phrased affected whether a student was able to respond correctly. For instance, more students were able to determine which event had the largest probability than which event had the smallest probability. The same thought process to determine probabilities is not
always applied in the same way, based on how questions are phrased. Other misconceptions have been found to be so prevalent that researchers have named them.

The *Outcome Orientation Misconception* is one in which students tend to make yes/no decisions based on a single event in a series rather than considering the entire series of events as a whole. A classic example of this misconception as described by Garfield and Chance (2000) is based upon a weatherman predicting rain. If the weatherman predicts a 70% chance of rain over the next 10 days, and it actually did rain on 7 of those 10 days, some students would argue that the weatherman did not do a good job of predicting the weather, because it should have rained on all of the days, since on any one given day, it is more likely to rain than not.

The *Representativeness Misconception* is one in which a student estimates the likelihood of an event based on how closely it resembles the population. Garfield (2002) gives an example of flipping a coin 6 times, and determining which outcome, HTHHTT or HTHHHH, is more likely. Although both outcomes are equally likely, students will often think the first of these outcomes is more likely since half of the flips were heads and half were tails. A second example of the representativeness misconception is the Gambler’s Fallacy (Garfield, 2002). After a long streak of heads while flipping a fair coin, students might think that tails is more likely to come up next, since it is “overdue.”

*Equiprobability Bias* is when all events tend to be viewed as equally likely despite the number of possible outcomes that comprise those events. Garfield (2002) described a classroom made up of an unequal number of business majors and science majors. When one person would be selected at random from this class, students would often incorrectly
assume that the probability of selecting a business major is the same as the probability of selecting a science major, since there were only two choices of major.

There are also a number of misconceptions that instructors have when it comes to assessing comprehension in their students. Instructors should recognize that correct calculations do not necessarily demonstrate an understanding of statistical ideas. As Rumsey (2002) described, being able to calculate a standard deviation does not demonstrate understanding of what the standard deviation is, how it helps describe a set of data, or how it might be used. Additionally, instructors often falsely believe that formulas help students to understand statistical ideas. Formulas might provide students an ability to obtain a result, but they do not help to explain how and when the formula should be used in a contextual situation (Rumsey, 2002). For instance, knowing the formulas for a permutation and combination does not help a student decide which formula was appropriate for a particular problem. A final misconception held by instructors is that when students explain solutions with statistical language, they are demonstrating an understanding of a statistical idea. Rumsey (2002) illustrated this point using hypothesis testing. When a student decides to reject the null hypothesis because the $p$-value is less than alpha, that same student might not understand what the rejection of the null hypothesis actually meant in terms of the problem or what the $p$-value to alpha comparison actually was doing. Keeping these misconceptions in mind, developing methods that truly assess student comprehension has been a recent area of research (Baharun & Porter, 2016; Garfield, 2003; Garfield & Ben-Zvi, 2007).
Assessing Comprehension

In 2003, Garfield described a need for new forms of assessment to evaluate the effectiveness of newly developed curricula and changes in the approaches to teaching. Garfield also suggested that new assessments were needed to explore the development of statistical reasoning skills in students (Garfield, 2003). As a result of the recent reform movement, many studies were conducted to determine the effectiveness of various interventions. A consistent problem in these quantitative studies was the lack of high-quality, consistent measures to assess the three types of student learning: statistical literacy, statistical reasoning, and statistical thinking (delMas et al., 2007; Garfield & Ben-Zvi, 2007). Most researchers had been using their own students at their own learning institutions and it had been a common practice to use their own final exams or course grades as a measurement of comprehension (Garfield & Ben-Zvi, 2007). While these provided researchers with personally meaningful results, their findings were not necessarily generalizable to the wider statistics education community due to the lack of reliability and validity in their outcome measures (Garfield & Ben-Zvi, 2007). Four attempts were made to develop assessments to address the lack of reliability and validity in the instructor-created assessments used in previous studies.

The Statistical Reasoning Assessment (SRA) was a first attempt to create a validated test that would evaluate the effectiveness of the new statistics curriculum in a high school setting (Garfield, 2003). At the time of its development, there were no other existing instruments that could assess conceptual understanding and statistical reasoning (Garfield, 2003). The SRA consists of 20 multiple choice questions and assesses reasoning about statistical measures, uncertainty, samples, and association. It also
assesses common misconceptions involving averages, outcome orientation, samples, representativeness, and equiprobability.

The Statistics Concept Inventory (SCI) was another attempt to develop a multiple choice, standardized test to measure understanding for a collegiate introductory statistics course (Allen, 2006). There are two versions of the SCI, a full exam and a shortened exam, both with acceptable levels of reliability and validity. However, it was found to be insufficient for some researchers, as the questions on this test were geared toward future engineering students (delMas et al., 2007).

A third set of questions was then developed as the Assessment Resource Tools for Improving Statistical Thinking (ARTIST) Project, funded by the National Science Foundation (Garfield et al., 2006). The ARTIST website provides an online database of over 1,000 questions, including both multiple choice and short answer formats, to assess statistical literacy, statistical reasoning, and statistical thinking (Garfield & delMas, 2010). Instructors can search for and select questions based on their choice of topic, question format, and cognitive outcome.

Out of the ARTIST Project, another assessment instrument, called the Comprehensive Assessment of Outcomes in Statistics (CAOS) test, was developed by a team of researchers to measure basic literacy and reasoning about descriptive statistics, probability, bivariate data, and basic statistical inference (Baharun & Porter, 2016; Garfield & delMas, 2010). Once created, it was amended four times based on instructor feedback, student feedback, and reliability testing (Baharun & Porter, 2016). In its final form, the CAOS test consists of 40 multiple choice questions and has acceptable levels of reliability and validity (delMas et al., 2007).
At the present time, there is not one unique existing instrument to assess comprehension of probability topics for students in an introductory statistics course. However, each of these assessments do contain some probability questions. The assessment developed for this dissertation sampled from many of the probability items found on these previously validated instruments.

**Connecting Comprehension and Attitude**

An increase in academic performance could be a positive influence on student attitude toward statistics. Di Martino and Zan (2010) found such a strong connection between perceived competence and emotional affect that the students essentially used them as synonyms. A well-established relationship exists between a positive attitude and higher achievement in statistics courses (Ramirez et al., 2012; Tempelaar et al., 2007). When students are presented with meaningful tasks in relevant context, they could experience higher levels of learning and achievement (Hay et al., 2015). Conversely, individuals who develop a negative attitude or a lack of interest are less likely to engage in that subject in the future (Carmichael et al., 2010). Students who experience frequent frustration in a content area may give up and avoid putting effort into understanding the material (Di Martino & Zan, 2010). It appears that student comprehension cannot be fully studied without also examining its interplay with attitude.

**A History of Attitude Toward Statistics**

The study of student attitude in statistics education has risen into prominence over the past quarter century. Gal et al. brought the concept of statistics attitude into the spotlight in their 1997 paper, *Monitoring Attitudes and Beliefs in Statistics Education*. Attitudes can be defined as “relatively stable, intense feelings that develop as repeated
positive or negative emotional responses are automatized over time” (Gal et al., 1997, p. 40). There is general agreement among researchers that attitude refers to the predisposition of an individual toward a content area. It is typically expressed in either positive or negative terms, yet caution should be taken toward reducing attitude to a simple binary outcome (Di Martino & Zan, 2010). Attitude is a multifaceted concept, and in studying the interaction of the underlying pieces, researchers have hoped to improve student outcomes in statistics (Di Martino & Zan, 2010; Emmioglu & Capa-Aydin, 2012; Vanhoof et al., 2011).

Perhaps the premiere measure for attitude toward statistics was developed by Candance Schau. The Survey of Attitudes Toward Statistics (SATS-36) expresses attitude in terms of six components: Affect, Cognitive Competence, Value, Difficulty, Interest, and Effort (Schau, 2003). This measure has been extensively validated and provides a solid foundation for analysis of attitude and its impact on student learning outcomes in statistics (Tempelaar et al., 2007). For the purpose of this research, attitude is considered to be constructed of these six components as defined by Schau (see Table 2.2).

Researchers have extensively documented the existing relationship between student attitude and achievement. By better understanding the extent of this relationship, educators can tailor instructional methods to better meet the educational needs of students (Emmioglu & Capa-Aydin, 2012).
Table 2.2

Six Attitude Components of the SATS-36 (Schau & Emmioglu, 2012, p. 87)

<table>
<thead>
<tr>
<th>Component</th>
<th>Definition</th>
<th>Sample Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affect</td>
<td>Students’ positive and negative feelings concerning statistics</td>
<td>“I am scared by statistics.”</td>
</tr>
<tr>
<td>Cognitive</td>
<td>Students’ attitudes about their intellectual knowledge and skills when applied to statistics</td>
<td>“I can learn statistics.”</td>
</tr>
<tr>
<td>Competence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value</td>
<td>Students’ attitudes about the usefulness, relevance, and worth of statistics in personal and professional life</td>
<td>“I use statistics in my everyday life.”</td>
</tr>
<tr>
<td>Difficulty</td>
<td>Students’ attitudes about the difficulty of statistics as a subject</td>
<td>“Most people have to learn a new way of thinking to do statistics.”</td>
</tr>
<tr>
<td>Interest</td>
<td>Students’ level of individual interest in statistics</td>
<td>“I am interested in using statistics.”</td>
</tr>
<tr>
<td>Effort</td>
<td>Amount of work the student expends to learn statistics</td>
<td>“I plan to work hard in my statistics course.”</td>
</tr>
</tbody>
</table>

The Influence of a Positive Attitude Toward Statistics

In a holistic sense, a positive attitude can lead to many desirable outcomes in statistics education. Long after students have forgotten the specific skills learned in a course, they will remember how they felt about the course (Ramirez et al., 2012). This general feeling, defined by Schau as affect, contains tremendous implications regarding statistics education, particularly in introductory statistics courses that serve as gatekeepers to future experiences in the content area. Student affect and interest toward a subject can influence their future course selections, which in turn plays into their likelihood to apply the content in the future (Hay et al., 2015).
In addition to affect, the attitude component of *cognitive competence*, or the belief by an individual in their own ability to learn statistics, has been found to play a significant role in student achievement in the classroom (Ramirez et al., 2012). Studies have shown that cognitive competence is a strong predictor of student achievement in statistics, even when compared to other high potential variables such as prior mathematical experience (Dempster & McCorry, 2009; Ramirez et al., 2012; Tempelaar et al., 2007). If students believe that they will find success in learning statistics, they are more likely to actually do so. As summarized by Schau and Emmioglu (2012), “A large number of educational and cognitive theories, with a great deal of research corroboration, support the conclusion that students’ attitudes toward the discipline are important course outcomes and, in fact, are at least as important as knowledge and skills in the discipline” (p. 86). In light of these potential benefits on student achievement, it is clear that a positive attitude is a worthy outcome in and of itself.

*Dangers of a Negative Attitude Toward Statistics*

While a great deal of research points to the influence of positive student attitudes on performance, an equal amount of support exists for the dangers of negative attitudes. First, it is important to understand that a negative attitude could refer to any of the interrelated components of the *SATS-36*, rather than a simple negative emotional affect (Di Martino & Zan, 2010). A host of undesirable outcomes are associated with negative attitude components. Students with a negative attitude are less likely to persevere through more courses in that content area (Gal et al., 1997). This becomes especially problematic when a negative attitude develops during an introductory statistics course. Studies have shown that many existing introductory statistics courses do not
improve student attitudes, or worse, have a negative effect on student outlook (Dempster & McCorry, 2009; Schau & Emmioglu, 2012). Students who exhibit lower interest in the content tend to learn less while actually in the course (Hay et al., 2015). Even worse, students who exit their first statistics course with a negative attitude are much less likely to ever apply what they learned in their future (Schau & Emmioglu, 2012).

Though it is important not to attribute a causal claim to the relationship between attitude and student achievement, it is undeniable that the two domains are connected (Di Martino & Zan, 2010). With such an important relationship established, it becomes crucial to learn how attitude can be changed. In other words, how can a negative attitude be transformed into one that is positive? Many students who have experienced a shift in attitude (both positive and negative) attribute it to a turning point sparked by either a particular instructor or course (Di Martino & Zan, 2010). Additionally, if students see more relevance and value in what they learn, there is evidence to suggest that they may experience an increase in cognitive competence and emotional affect (Di Martino & Zan, 2010). Under the existing collegiate introductory statistics structure, the majority of students enter the course with a neutral attitude toward statistics (Schau & Emmioglu, 2012). Unfortunately, student attitude seldom improves over the course of the class, and sometimes even becomes more negative (Schau & Emmioglu, 2012). It is evident that something more must be done to improve student outlook toward the content area. The addition of relevant context has the potential to be a solution to this problem.

**Other Influences on Student Learning Outcomes**

Existing research has examined many other factors to determine whether they have an effect on student comprehension and attitude. Of all such factors, qualities of the
teacher and the choices made within their instruction represent perhaps the strongest influence on student achievement (Hattie, 2015). The support provided by a teacher and the relationship that they form with students can have a major effect on student engagement and attitude toward mathematics (Davadas & Lay, 2018; Di Martino & Zan, 2010; Hay et al., 2015). By keeping students engaged even as the content gets more challenging, teachers can guide students toward deeper levels of interest in the subject matter (Hidi & Renninger, 2006).

In addition to the impact of the instructor, the length of exposure to a content area can play a role in resulting attitude and comprehension for students. Attitude toward mathematics can be somewhat resistant to change (Di Martino & Zan, 2010), and enduring shifts in attitude must be reinforced over time (Gal et al., 1997). Deeper levels of interest in a subject are achieved through sustained contact with the material (Hidi & Renninger, 2006). In order for a significant change in attitude to occur, the intervention under question must occur for long enough to reinforce the desired outcome.

Finally, mixed findings exist regarding the relationship between student age and attitude toward a subject area. In middle school, students have been observed to have a more negative attitude as they age (Carmichael et al., 2010). Yet when considering the population of college students, those who were older exhibited higher levels of interest toward statistics (Bateiha et al., 2020). In general, the relationship between age and student attitude is too contradictory to warrant sweeping conclusions on its impact (Ramirez et al., 2012). However, research does suggest that early positive exposure toward a content area makes a significant difference in eventual attitude toward that subject (Hidi & Renninger, 2006). Thus, the setting of middle school, high school, or...
college must at least be considered when attempting to improve student attitude toward statistics.

Summary

In its current form, statistics education is not reaching its full potential for student achievement. Lyford et al. (2019) suggested, “students often balk at the dreary context of flipping a coin or drawing black and white balls from urns…” Innovative strategies must be incorporated to promote student comprehension and to improve student attitude toward statistics. Including relevant context as an intervention into the instructional strategy is one possible solution. Situated Learning Theory and Realistic Mathematics Education both provide a framework for the development of authentic, situated learning experiences centered around the use of relevant context. Through the use of a mixed methods approach, this study sought to determine the effect of relevant context on comprehension and attitude within the probability unit of a collegiate level introductory statistics course.
Chapter 3

Methodology

Appropriateness of the Research Design

In order to determine the extent to which student comprehension and attitude toward statistics are impacted by the inclusion of relevant context in lessons, a mixed methods, explanatory design was used. For the initial quantitative phase, a quasi-experimental design was implemented. Participants could not be randomly assigned to their course section, making a true experimental design impossible. A quasi-experimental design is frequently used in educational research, although it poses additional threats to validity due to the potential differences between treatment groups (Creswell & Guetterman, 2019). To account for variability in individual student performance, the instrumentation for both comprehension and attitude incorporated a pretest/posttest approach.

In addition to the quantitative analysis, a qualitative follow-up was conducted to better understand the experience of participants during the intervention. Qualitative research “represents the meanings given to events by the people who live them, not the values, preconceptions, or meanings held by the researchers” (Yin, 2016, p. 9). To minimize the bias in qualitative research, it is important to establish strategies to validate the methodology and findings including peer reviews, audits, and relating the work to previous research (Yin, 2016). Qualitative methods which included triangulation were employed to investigate the impact of using relevant context on student experience in the course.
Research Questions and Hypotheses

Using the comprehension assessments, attitude surveys, and interview or focus group responses, an attempt was made to answer the following three research questions:

1. To what extent does using relevant context within situated experiences affect student comprehension of probability concepts?
2. To what extent does using relevant context within situated experiences affect student attitude toward statistics?
3. In what ways do students describe the factors that affected their comprehension of probability and their attitude toward statistics based on the probability unit?

Depending on the course section selected during registration, students became part of either the treatment group or the control group.

- Students in the two sections comprising the treatment group engaged in active learning activities centered around relevant context that tied into student interests, as identified through a preliminary interest survey and verified through exit slip surveys.
- Students in the one section comprising the control group engaged in active learning activities centered around context that was not tailored to student interests.

With these two groups in mind, the following hypotheses were tested:

RQ#1  \( H_{01} \): There is no significant difference in average comprehension gains between students in the treatment and control groups.
H_{A1}: There is a significant difference in average comprehension gains between students in the treatment and control groups.

RQ#2 \ H_{02}: There is no significant difference in any of the six attitude components of Affect, Cognitive Competence, Value, Difficulty, Interest, or Effort between students in the treatment group and students in the control group.

H_{A2}: There is a significant difference in at least one of the six attitude components of Affect, Cognitive Competence, Value, Difficulty, Interest, or Effort between students in the treatment group and students in the control group.

Because the third research question is qualitative, no assumptions were made about the student responses and therefore, no hypotheses were tested.

**Research Design**

A quasi-experimental, mixed methods, explanatory approach was used to examine the impact of relevant context on student learning outcomes. Changes in student comprehension and attitude were measured quantitatively using pre-intervention and post-intervention assessments. Additionally, a thematic understanding of the ways students describe their change in comprehension and attitude was examined through a qualitative lens using individual interviews and semi-structured focus groups. A case-study approach was provided for the qualitative portion of the research. An overview of the research design is provided in Figure 3.1.
Figure 3.1

Overview of Research Design

<table>
<thead>
<tr>
<th>Planning and Quantitative Pre</th>
<th>Intervention Activity Design</th>
<th>Intervention Period and Quantitative Post</th>
<th>Qualitative Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot Studies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevant Context Identification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre SATS-36 Attitude</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention Activity Design</td>
<td>Video Lectures / Homework (Not modified for study)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post SATS-36 Attitude</td>
<td>Posttest Comprehension</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exit Slip Assessing Relevance after each Activity</td>
<td>Posttest Comprehension</td>
<td></td>
<td>Focus Groups</td>
</tr>
</tbody>
</table>

Treatment and Control Group

Three sections of an introductory statistics course at an urban midwestern university were available for use during the semester under study. The intervention took place over a six-week unit covering probability and probability distributions. Two sections were selected to receive the relevant context-based intervention, and the third section served as a control group, participating in similar lessons and activities that incorporated other contexts not tailored to student interest. Due to necessity, all three sections were taught by different instructors. A professor with 17 years of experience who is also a member of the research team taught one section of the treatment group, Treatment Group 1. The second section of the treatment group, Treatment Group 2, was taught by an adjunct professor with three years of experience, and the control group was taught by a graduate student with three years of experience. Students in all three sections received instruction via the same set of videos created by the professor who was assigned Treatment Group 1. All three sections participated in engaging activities during class time that were intended to reinforce the concepts introduced in the instructional videos. Differences between the structure of the three sections are summarized in Table 3.1.
Table 3.1

Characteristics of Treatment Groups and Control Group

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Treatment Group 1</th>
<th>Treatment Group 2</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days/week course met</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Days/week of intervention</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Instructor</td>
<td>Professor / Researcher</td>
<td>Adjunct professor</td>
<td>Graduate student</td>
</tr>
<tr>
<td>Instructor experience</td>
<td>17 years</td>
<td>3 years</td>
<td>3 years</td>
</tr>
<tr>
<td>In-class activity</td>
<td>Relevant context</td>
<td>Relevant context</td>
<td>Other context</td>
</tr>
</tbody>
</table>

Relevant Context Identification

Before the beginning of the six-week probability unit under study, students in all three sections of the course were given an interest inventory created for the study (see Appendix B.1). Results from the interest inventory were used to identify potential contexts for intervention activities that students found most relevant, such as “Sports” and “TV Shows” and used to reject potential generic contexts for the control group activities. Contexts with higher degrees of relevance were marked as potential settings for the intervention activities. Contexts with lower degrees of relevance were eliminated from the potential activities for the treatment group. The full results of the interest inventory can be found in Appendix B.2, and a summary of the contexts selected is provided in Chapter 4.

Intervention Activity Design

One of the potential contexts that paired well with the content of the lesson was selected, and an activity was designed which merged content with relevant context. Design of the activities was guided by the theoretical framework which merged Situated
Learning Theory and Realistic Mathematics Education. Both of these theories promote the exploration of authentic real-world experiences in the acquisition of new knowledge (Van den Heuvel-Panhuizen, 2020; Zakarita & Syamaun, 2017). Ten total activities were created for the treatment group which tied the probability unit objectives to a context that the majority of students identified as personally relevant. This relevance sought to incorporate a deep level of contextualization to promote engagement and buy-in from students (Clinton & Walkington, 2019). Ten parallel activities were then created and were centered around contexts which were either included in the interest inventory and identified as not particularly relevant by a majority of students or not included in the interest inventory and believed to be not highly relevant. The corresponding activities were then presented to students in the treatment and control groups over the course of the unit. After each lesson, students completed a short exit slip which served both as an attendance record and also allowed students to rate the personal relevance of the context used in the daily activity.

**Quantitative Design**

Prior to the intervention unit, students completed a pre-version of the *Survey of Attitudes Toward Statistics (SATS-36)* attitude survey, which breaks attitude into the six components described in the hypotheses. Students also completed a pretest assessing their comprehension of probability, which contained a combination of items from published, validated statistics exams along with instructor-created items (see Appendix C.1). At the end of the unit, but prior to the posttest, students in all three sections completed the post-version of the *SATS-36* (see Appendix C.3), which enabled researchers to measure any change in individual attitudes toward statistics. Students then completed the posttest (see
Appendix C.2) to measure their comprehension of probability objectives covered during the unit. This posttest contained the same 10 items from the pretest, along with additional questions to ensure all unit objectives were assessed.

**Qualitative Design**

At the beginning of the intervention, individuals from each of the three sections indicated their interest in sharing their experiences in either a one-on-one interview or a focus group by completing the Volunteer / Extra Credit Opportunity form (see Appendix A.2). In exchange for their participation, selected individuals received extra credit points in the course. Students were deemed eligible for participation in an interview or focus group if they had attended at least six out of the 10 lessons in the intervention. Selection was also based on student preference for interview or focus group and on student availability to meet during the interview period. Students who were not selected for the qualitative portion of the study were also eligible to receive extra credit by responding to prompts similar to interview questions in a supplementary written assignment. These written assignments were not included as a part of the study since there was no opportunity to ask follow-up questions.

The interviews were conducted within a two-week period after the students completed the comprehensive posttest and the post-version of SATS-36 attitude survey. All qualitative interviews and focus groups were conducted over Zoom, using a semi-structured approach, where students were able to describe factors affecting comprehension and attitude over the probability unit. Appendix C.4 contains the interview questions used during the qualitative portion of the study. Interviews and focus
groups were conducted by a single member of the research team, who was not involved in instruction of the course.

**Pilot Studies**

In the semester prior to the study, researchers piloted a draft version of the interest inventory. Using these preliminary student responses to the survey, a more detailed interest inventory was developed that spanned nearly 50 unique potential contexts for class activities. Changes to formatting were also made in an effort to improve survey clarity.

An eight-question draft version of the pretest was piloted in all sections of the course at the university during the semester prior to the study. All questions on the piloted pretest were taken directly from the Comprehensive Assessment of Outcomes in a first Statistics course (CAOS) exam, the Assessment Resource Tools for Improving Statistical Thinking (ARTIST) test bank, and the Statistical Reasoning Assessment (SRA). The draft pretest, given the week prior to the probability unit, aided in assessing the clarity of the questions and evaluating the prior knowledge in a group similar to the one to be used in the study. Based on the results of the pilot, an additional distractor answer was added to some multiple-choice items so that all pretest items had the same number of incorrect options. Additionally, the types of questions where students were instructed to “select all that apply” were eliminated so that each pretest item had only one correct answer. Also, questions that the majority of students were able to answer correctly without any exposure to the probability unit were eliminated from the pretest. Finally, the total number of questions was increased from eight to ten so that more course topics could be assessed.
Additionally, interview questions used during the qualitative portion of the study were piloted for an assignment within the doctoral program in which three individual interviews were conducted using student volunteers who were familiar with the use of relevant context in statistics. By conducting these preliminary interviews, interview questions were modified to provide additional clarity and elicit greater detail in responses.

**Setting and Participants**

The convenience sample was composed of 72 students in an urban midwestern university taking Introductory Probability and Statistics during the Spring 2022 semester. Participants were selected based on enrollment in the three face-to-face sections of the course. All students who enrolled in one of these sections during Spring 2022 and remained enrolled throughout the duration of the study became eligible for participation. Demographic data was collected on all students as a normal part of the coursework during the first week of the semester. Out of the initial 72 students in the sample, 43 were members of the treatment group and the remaining 29 were part of the control group. The treatment and control groups were comparable in the demographic categories of educational level (Figure 3.2), gender (Figure 3.3), citizenship (Figure 3.4), and choice of major (Figure 3.5). Further demographic data can be found in Appendix E.1.
Figure 3.2

*Educational Level Comparison*

![Educational Level Comparison](image)

Figure 3.3

*Gender Comparison*

![Gender Comparison](image)

Figure 3.4

*Citizenship Comparison*

![Citizenship Comparison](image)
A total of 16 students, 22% of the sample, participated in the qualitative portion of the study. At the beginning of the intervention, students identified whether or not they were willing to participate in a one-on-one interview or focus group. Ten one-on-one interviews were conducted: three students from Treatment Group 1, four students from Treatment Group 2, and four students from the control group. Two focus groups were conducted. Three students from Treatment Group 1 formed one focus group and two students from the control group formed the other. Due to student scheduling conflicts, it was impossible to conduct a focus group for Treatment Group 2. The number of students participating in the interviews and focus groups are summarized below in Table 3.2.

**Table 3.2**

*Number of Students Participating in Interviews and Focus Groups*

<table>
<thead>
<tr>
<th></th>
<th>Treatment Group 1</th>
<th>Treatment Group 2</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three one-on-one</td>
<td>Four one-on-one interviews</td>
<td>Four one-on-one interviews</td>
<td></td>
</tr>
<tr>
<td>interviews</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One focus group with three students</td>
<td>One focus group with two students</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Instrumentation

Relevance

To initially identify contexts that participants would find relevant, a Student Interest Inventory (see Appendix B.1) was created. This survey incorporated a five-point Likert scale that ranged from $1 = \text{Extremely Uninterested}$ to $5 = \text{Extremely Interested}$. The survey first asked participants to rate their interest in experiencing in-class examples centered on seven broad categories: “Sports & Activities,” “Health, School, & Career,” “Politics & Social Justice,” “Leisure Activities & Hobbies,” “Technology,” and “Finance & Business.” Students rated their interest on a total of 49 specific issues, each under the umbrella of one of the seven broad categories. For example, within the “Finance & Business” section, students were able to rate their interest in lessons centered on the stock market, housing prices, insurance rates, and several other specific contexts.

During the intervention, a daily exit slip (see Appendix B.3) created for the study served two purposes. First, it asked students to rate the personal relevance of the activity for the day on a scale of $1 = \text{Not at all Relevant}$ to $5 = \text{Highly Relevant}$. This served as evidence of whether the lessons created for the treatment group actually achieved the desired relevance and whether the lessons created for the control group were perceived as non-relevant. Second, the exit slips also represented a way to keep track of student attendance. Based on exit slip completion, researchers were able to record how many of the 10 intervention lessons were attended by each student in the study (see Appendix E.2).
Quantitative - Comprehension

Student comprehension of probability was assessed through a pretest and posttest which had been validated by a panel of content experts at the university under study. These tests included instructor-created items along with validated items from the CAOS exam, the ARTIST website, and the SRA.

Change in student comprehension of probability concepts was measured quantitatively through one pretest and a posttest which was split into two parts. To create these tests, a small panel of experts in the field of probability and statistics were consulted. This panel was composed of six collegiate-level instructors who had all taught a collegiate level statistics course at the urban midwestern university under study and all of which held at least a master’s degree in mathematics or statistics. The panel rated all questions on the pretest and posttest for content validity using a 5-point Likert scale. Questions were rated both for clarity and phrasing and for how well the question addressed the intended concept to be assessed. All but one of these questions received an average rating greater than 4 (see Appendix C.2). The one question that received the lowest rating (3.2) was taken directly from the CAOS test. All items on the CAOS test had previously been validated by a larger panel of experts (delMas et al., 2007). As such, it was determined that it would be appropriate to use to measure comprehension.

The pretest consisted of 10 questions; six were multiple choice and the remainder were computational questions with numeric solutions. Half of the questions on the pretest were selected from the CAOS test, the ARTIST test bank, and the SRA, all of which have gone through analysis and have appropriate levels of validity and reliability (delMas et al., 2007; Garfield, 2003). Due to a lack of previously validated instruments measuring the
specific probability concepts for the course under study, the remaining five questions on the pretest were authored by one of the researchers. This ensured that the content in the assessment was representative of the entire probability unit. Four of the 10 questions addressed statistical literacy, while the remaining six addressed statistical reasoning.

The entire posttest included these same 10 pretest questions along with 18 other exam questions, some having multiple parts. The CAOS, ARTIST, and SRA do not have questions that address all concepts for the course used. Therefore, these additional 18 questions were either identical to or adaptations of questions from CAOS, ARTIST, and SRA, or they were written by the researchers. The posttest was specifically designed this way to ensure all topics taught in the probability unit were assessed. Historically in this introductory course, the unit assessment on probability is split into two parts, and this format was chosen to continue with this precedent. Therefore, the posttest had two parts, corresponding to the two modules students use during the probability unit. Five questions from the pretest appeared on the first half of the posttest and the remaining five appeared on the second half of the posttest. Students had the option to study for and take the two parts of the posttest on separate days.

The first half of the posttest contained 11 questions pertaining to basic probability concepts, probability formulas, contingency tables, and counting rules. Slightly under 50% of the questions on this portion of the posttest assessed statistical literacy and slightly over 50% assessed statistical reasoning. The second half of the posttest contained 17 questions pertaining to discrete and continuous probability distributions. Slightly more than 50% of the questions on this portion of the posttest assessed statistical literacy and slightly under 50% assessed statistical reasoning (see Table 3.3).
The same panel of experts from the university in which students were enrolled also rated both halves of the posttest for face validity (see Appendix C.2) on a 5-point Likert scale. The first half received an average score of 5/5 and the second half received an average score of 4.6/5. Therefore, it was concluded that the posttest appropriately assessed student comprehension of the probability unit.

**Table 3.3**

*Number of Questions Addressing Each Component of Comprehension on Assessments*

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Questions for Literacy</th>
<th>Questions for Reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>4 (40%)</td>
<td>6 (60%)</td>
</tr>
<tr>
<td>Posttest (Part 1)</td>
<td>5 (45%)</td>
<td>6 (55%)</td>
</tr>
<tr>
<td>Posttest (Part 2)</td>
<td>9.75 (54%)</td>
<td>7.25 (46%)</td>
</tr>
<tr>
<td>Posttest (Overall)</td>
<td>14.75 (52%)</td>
<td>13.25 (48%)</td>
</tr>
</tbody>
</table>

**Quantitative - Attitude**

Change in attitude regarding statistics was measured quantitatively through a pre-version and post-version of the *SATS-36* developed by Candace Schau. The *SATS-36* breaks attitude into six distinct components (Schau & Emmioglu, 2012):

- Affect
- Cognitive Competence
- Value
- Difficulty
- Interest
- Effort

The *SATS-36* is one of the best known and most validated measures for attitude toward statistics (Emmioglu & Capa-Aydin, 2012; Tempelaar et al., 2007). It has “good
to excellent psychometric properties” including “moderate to high values for internal consistencies for each of the attitude components evaluated using Cronbach’s coefficient alpha,” as well as high concurrent validity (Ramirez et al., 2012, p. 61). Extensive research has been conducted on the individual questions within the SATS-36, on whether any defy the pattern of the others or whether any unnecessary redundancies exist (Vanhoof et al., 2011). Use of the SATS-36 is widespread across the world (Schau & Emmioglu, 2012), and the use of this measure can add to the existing knowledge base on statistics attitude.

Other well-known instruments for this purpose include Roberts & Bilderback’s Statistics Attitude Survey and Wise’s Attitudes Toward Statistics Scale. However, these instruments only measure one and two dimensions of attitude respectively (Bechrakis et al., 2011). It is widely accepted that attitude is multi-dimensional by nature (Di Martino & Zan, 2010), and the aforementioned measures have diminished in popularity in recent years (Bateiha et al., 2020), likely due in part to their lack of dimensions.

In analyzing student results on the SATS-36, it is not appropriate to create a single composite score for attitude; each of the six components must be considered separately (C. Schau, personal communication, May 22, 2021; Di Martino & Zan, 2010). Each of the 36 questions on the survey is rated on a 7-point Likert scale and is tied to one of the six attitude components listed above (Schau, 2003). A higher score on a question corresponds to a more positive attitude and scoring on negatively worded questions is reversed. For questions within each of the six components, the corresponding scores are then averaged to create six attitude component subscores (Schau, 2003).
Qualitative - Comprehension & Attitude

Qualitative research was chosen to answer the third research question concerning the ways students describe their change in comprehension of probability and their attitude toward statistics, based on the use of relevant context. The overarching question for the qualitative research was to determine in what ways the students described the factors that affected their comprehension of probability and their attitude toward statistics based on the probability unit. As individuals react differently to one-on-one interviews and focus groups, both options were provided as part of the qualitative research. The semi-structured format allowed for the same questions during the interviews and focus groups with open ended responses and additional questions included, as necessary.

The interview questions (see Appendix C.4) were developed in order to provide further insight about results from the initial two research questions. The questions were designed to be open-ended in order to create an environment where participants felt comfortable conversing and they would not feel as if they were being studied (Yin, 2016). The questions were separated into sections allowing the students to focus on comprehension and then attitude. The initial question for each section asked students for a self-assessment of their comprehension of probability topics and their attitude toward statistics. The next question asked the students to describe their change in comprehension and attitude over the unit. The students were asked about elements of the instructional process and which of those were most crucial to their comprehension and attitude during the probability unit. To gain a direct understanding about the use of the context, a question was asked about how context used during the activities affected their comprehension and their attitude. Additionally, the students were asked whether any
outside circumstances such as sickness, family circumstances, or work affected their comprehension and attitude. This outline and method of questioning provided an opportunity for students to describe their change, if any, in both comprehension and attitude.

During the qualitative pilot study, a trial run of the interview questions was conducted with three student volunteers who were already familiar with the use of relevant context in statistics lessons. These trial interviews provided the opportunity to confirm clarity of wording and adjust questions for effectiveness and efficiency. The trial run allowed the researchers to pilot the interview script and list of questions. The original questions changed and evolved with each subsequent interview. The initial questions did not dive very deep into discovering the core thoughts on how relevant context impacted the comprehension and attitude of participants. For example, the interview question, “How did the context used during the in-class activities affect your comprehension / attitude toward statistics?” was added to inquire as to their insight on the use of context. In addition, a question was added about the effect of outside circumstances since that was mentioned in all three trial interviews as influencing comprehension and attitude.

**Procedures for Data Collection**

**Relevance**

Students completed the interest inventory before the start of the probability unit where the intervention took place. At the end of each in-class activity, students completed a short exit slip rating the personal relevance of the context for the day. The inventory was completed online using the class webpage, and exit slips were completed on paper at the end of each class.
**Quantitative - Attitude**

Prior to the start of the intervention unit, students completed the pre-version of the SATS-36 as an assignment on the class webpage. This yielded a baseline for each of the six attitude components for each student, expressed as an average between 1 and 7. Upon completion of the unit, but before the posttest was given so as to not allow test performance to influence attitude, students were given the post-version of the SATS-36 on the same webpage. Questions on the post-version of the survey were identical, except for a switch to past tense from the future tense used on the pre-version (Schau, 2003).

**Quantitative - Comprehension**

Before the probability unit began, students were given a 10-question pretest administered outside of class in an online format, assessing their understanding of core probability concepts. These same 10 questions were embedded into the posttest, and the difference in performance for these specific questions (postscore - prescore) was calculated for each student. In addition to this data, each student also earned a score from 0 to 100 for their performance on the entire end-of-unit posttest. The posttest was also administered outside of class using an online format with automated test proctoring software provided by the university. Students were allowed to complete both parts of the exam at their convenience, but no later than six days after the final lesson had been delivered.

**Qualitative - Comprehension & Attitude**

Following the unit assessment, volunteers from both the control and treatment groups were selected for participation in individual interviews or separate focus group sessions. Volunteers were eligible based on their selection in the Volunteer / Extra Credit
(see Appendix A.2) form and having attended more than half of the classes in the intervention unit. Using a semi-structured approach, interview questions were posed to analyze the ways in which students described their change in comprehension of probability and their attitude toward statistics, based on the use of relevant context. The focus groups and interviews of the qualitative analysis were conducted via Zoom. The one-on-one interviews lasted between 20 and 35 minutes with the focus groups lasting between 45 and 60 minutes. A standard script was used in conducting the interviews for consistency purposes, but the semi-structured format allowed the interviewer to probe for further information as needed. Each of the Zoom sessions were recorded to assist with transcribing and coding. Anonymity of the participants was protected, as recordings were stored on a secure password-protected computer, and all identifiable participant information was removed from the transcripts.

**Data Processing and Analysis**

**Relevance**

A list of the topics was compiled and identified as most relevant to student interest after students completed the interest inventory (see Appendix B.2). The lessons were primarily focused on topics which a majority of students rated the context as either “Somewhat” or “Extremely” Interesting. Activities were then developed and centered around 10 of the most popular topics, giving priority to those with the strongest connection to the probability content.

For the daily exit slips, each student recorded a score between 1 and 5, based on their perception of the personal relevance of that activity. The mean relevance score was then calculated for each lesson in the treatment group and in the control group by
averaging the ratings of each student in the group for that day. Additionally, the mean relevance rating for each individual student was calculated by averaging their scores for all 10 of the lessons (see Appendix B.3). This served as the basis for how well the activities of the intervention resonated with each individual student. All relevance rating data was analyzed using Excel, and graphics created to display relevance were also created through the use of Excel.

**Quantitative - Comprehension**

Each student received an improvement score (postscore - prescore) on a scale from -10 to 10 points, based on their change in performance on the 10 questions first given on the pretest. These improvement scores were then averaged for all students in the treatment group, creating an overall average improvement score. The average improvement score was calculated similarly for all students in the control group. Then, a two-sample t-test was conducted to determine whether the difference in average improvement was statistically significant between the two groups.

Additionally, each student received a score out of 100 on the end-of-unit test. These scores were averaged for all students in the treatment group, as well as for those in the control group. The averages were again compared using a two-sample t-test to determine whether a statistically significant difference existed between the average group performances. All statistical analyses and graphical displays for comprehension were created using SAS statistical software.

**Quantitative - Attitude**

For each student, the change in each of the six attitude components was calculated by finding the difference between the post-average and pre-average. The change for each
student in the treatment groups was then averaged for each of the six attitude components, creating six different averages for the treatment group. These six scores represent the average change in each component of attitude for students who experienced the relevant context intervention. The corresponding six averages were also calculated for students in the control group. Due to the relatively small sample sizes and the use of ordinal Likert data, the distribution of treatment and control data was analyzed using the non-parametric Mann-Whitney U test. Evidence suggests that the Mann-Whitney U test possesses higher statistical power than the corresponding parametric t-test on 7-point Likert data (Nanna & Sawilowsky, 1998). Other analyses find the power between these tests on Likert data to be very similar, with a slight edge given to the Mann-Whitney U test in situations where data is strongly skewed (deWinter & Dodou, 2010).

In addition to the comparison of attitude distribution, results were analyzed in a second manner consistent with prior research by the designer of the SATS-36. According to Candace Schau, the primary creator of the SATS-36, a difference of a half point or more in average should be considered significant (Schau & Emmioglu, 2012). A special note was taken for any student who changed in attitude by more than 0.5. The percentage of students in each group who changed by at least this amount was calculated for each of the six attitude components. Attitude data was analyzed and displayed graphically using SAS statistical software, with the exception of the ½ point analysis described above. This additional analysis was conducted using Excel.

**Qualitative - Comprehension & Attitude**

A single researcher conducted both the interviews and focus groups. The interview and focus group recordings were transcribed initially by Zoom with corrections
made, as necessary, by the researcher conducting the interview. Pseudonyms were used for each of the students to protect their identity. Transcripts were uploaded into a qualitative data analysis software, NVivo. The software allowed the researchers to group relationships in the qualitative data, assign codes to the transcripts, and begin to translate the data into themes.

Since research questions were previously developed based on the established theoretical framework, a deductive approach was used to analyze the qualitative data. A deductive approach provides an opportunity to start with the initial theories (Yin, 2016). Yin (2016) describes deductive coding as a process where categories are first identified, and they then lead to identification of relevant data. The responses were organized as confirmations of the hypotheses developed for each research question with a deductive approach helping “to establish the importance of a study” (Yin, 2016, p. 101).

Mayring (2000) provides a detailed step-by-step model to apply deductive category coding to data (see Figure 3.6). Their approach includes reliability and validity checks. For trustworthiness, a triangulation strategy was employed. This strategy included collecting data through the qualitative surveys, previous research, and the quantitative interviews.

An iterative process was used to analyze and interpret the data prior to defining and sorting the resulting codes into categories. Themes in participant responses emerged. Through coding, further information was revealed regarding the factors that affect student comprehension and student attitude.
Ethical Considerations

All subjects who participated were at least 18 years of age and gave their informed consent (see Appendix A.1) before participating in any research that fell beyond the routine structure of the course. Although all participants in the study were exposed to the same course content, only students in the treatment group had access to the relevant context intervention. Any personal participant information was only visible to the coordinator of the course, who is a member of the research team. Personal data was stripped of all identifying information before being analyzed. Each student was assigned a random number by the course coordinator, so that student confidentiality was protected even from other members of the research team. Transfer of student information between the course coordinator and other members of the research team was done in a manner that
kept student confidentiality intact. Results published within this paper are only shown in aggregate and cannot be traced back to individual students.

Upon completion of the intervention unit, students that expressed interest in participating in an individual interview or a focus group were contacted to share the schedule and to ascertain if they were still wanting to participate. Student participation in follow-up interviews and focus groups was strictly voluntary, with the ability to withdraw from participation at any time. Any data generated by these interviews was kept entirely confidential, with all recordings and transcripts stored on password-protected computers and all identifying information removed. Participant names were coded independently by the interviewer. The course coordinator, course instructors and other research team members did not have access to this information. All recordings were destroyed upon completion of the study. No known harm or mental anguish was caused to participants based on participation in the study.

**Summary**

A quasi-experimental, explanatory mixed-methods design was used to determine the impact that relevant contextual examples had on student comprehension and attitude toward statistics over a six-week unit on probability. Students who enrolled in one or the other of two sections of an introductory probability and statistics course were used as the treatment group, while students who enrolled in the third section of the course were used as the control group. All students received similar in-class activities which highlighted the major concepts in the unit. Activities for the treatment group were tailored to contain contexts that students had previously identified as potentially personally relevant, while activities for the control group used settings that were believed not to be particularly
relevant to students. Relevance ratings for each lesson were determined based on the exit slip data. Changes in student comprehension were measured quantitatively using $t$-tests based on the average change in pretest and posttest results and the overall average posttest scores, while changes in attitude were analyzed with the Mann-Whitney U test. Further information was gathered qualitatively at the conclusion of the probability unit through the use of semi-structured interviews and focus groups. Themes emerged from this information to provide further knowledge on factors that affected both comprehension and attitude.
Chapter 4

Results / Findings

Introduction

Using an explanatory mixed-methods approach, an intervention incorporating relevant context was tested in the probability unit of an introductory college statistics course. A quasi-experimental design was used, involving 72 students enrolled in three face-to-face sections of an introductory statistics course at an urban midwestern university. Two of the sections comprised a treatment group that experienced in-class activities centered around relevant context, as identified by an initial interest inventory. The remaining section of the course was used as a control group. Students in this section participated in similar in-class activities as the treatment group, but the contexts of these activities were not tailored to student interests. Participants in the treatment and control groups were comparable in gender, educational level, citizenship, and choice of major, although all three sections of the course were taught by different instructors.

Prior to the start of the intervention, students completed a probability pretest and the pre-version of the Survey of Attitude Toward Statistics (SATS-36). After each class during the six-week intervention, students completed an exit slip where they rated the contextual relevance of the daily activity. Results of the exit slips were used to assess whether activities completed by the treatment group really were more relevant than those completed by the control group. Upon completion of the unit, students were given the post-version of the SATS-36, followed by the unit exam, to determine whether they experienced any shift in their attitude or comprehension toward statistics or probability. Following the unit exam, 16 students were selected to participate in either a one-on-one
interview or a small focus group. Through these two types of interviews, insight was gained into driving influences on student comprehension and attitude over the course of the unit. Analysis of the quantitative and qualitative data was guided by three research questions.

**Research Questions**

1. To what extent does using relevant context within situated experiences affect student comprehension of probability concepts?
2. To what extent does using relevant context within situated experiences affect student attitude toward statistics?
3. In what ways do students describe the factors that affected their comprehension of probability and their attitude toward statistics based on the probability unit?

**Quantitative Results**

**Interest Survey**

**Data Collection.** During the first week of the semester, prior to the probability unit under study, students were asked to fill out an interest survey (see Appendix B.1) so that the personal interests of treatment group students could be incorporated into the contexts used for their in-class activities. Topics with high relevance ratings were avoided for the in-class activities in the control group, although the lowest rated contexts were not deliberately selected for control lessons. Students in both the treatment and control groups were asked to rate 49 potential lesson contexts on a five-point Likert scale ranging from *Extremely Uninterested* to *Extremely Interested*. The interest inventory was completely anonymous, but students were asked to identify which of the three course sections they were enrolled.
**Intervention Fidelity.** Even though there were only 72 students enrolled in the three course sections, there were 83 responses to the interest survey. Therefore, some students submitted responses more than once. However, because of the anonymous responses, it was impossible to eliminate duplicate results from individual students. Since this survey was only meant to guide the researchers in their planning of contexts and the exit slips would verify the amount of relevance for each particular context, the duplicate responses were not considered problematic.

**Results.** Based on responses to the interest inventory, 10 contexts were identified that had the potential to be found relevant by the majority of students within the treatment groups. After these topics were selected and in-class activities were created, 10 parallel lessons were created for the control group that centered around contexts not identified as highly relevant by students. Full results from the interest inventory can be viewed in Appendix B.2. Information on the 10 contexts selected for the study can be found in Table 4.1. All activities can be found in Appendix D.1 and Appendix D.2.

**Exit Slips**

**Data Collection.** Students received exit slips at the end of each activity and were asked to complete them before leaving the classroom (see Appendix B.3). Each exit slip contained the definition of the words “context” and “relevant” so that students could refer to these definitions in filling out their response. The instructor of the class also told the students what specific context they were to be rating on the exit slip after their activity. Students rated the context using a five-point Likert scale, with 1 representing *Not at all relevant* and 5 representing *Highly relevant*. The exit slips measured attendance as well as determined if each student found the context of the activity personally relevant.
### Table 4.1

*Contexts Chosen for the Situated Learning Activities as They Relate to the Interest Survey*

<table>
<thead>
<tr>
<th>Activity</th>
<th>Content</th>
<th>Context Category &amp; Subtopic from Survey for Treatment Group</th>
<th>Interest Survey Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Basic Probability and Law of Large Numbers</td>
<td>Technology - TV Binge Watching / Screen Time</td>
<td>86.7%</td>
</tr>
<tr>
<td>2</td>
<td>Addition Law</td>
<td>Leisure Activities &amp; Hobbies - Travel</td>
<td>92.8%</td>
</tr>
<tr>
<td>3</td>
<td>Tree Diagrams and Multistage Probability</td>
<td>Health - Contraception / STIs / Sex Issues</td>
<td>80.7%</td>
</tr>
<tr>
<td>4</td>
<td>Conditional Probability</td>
<td>Health - Psychology / Social Behaviors; Leisure Activities &amp; Hobbies - Pets</td>
<td>Psychology - 89.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pets - 86.7%</td>
</tr>
<tr>
<td>5</td>
<td>Discrete Random Variables and Expected Value</td>
<td>Health - Healthcare / Medicine; COVID-19 / Pandemics</td>
<td>Healthcare - 85.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>COVID - 68.7%</td>
</tr>
<tr>
<td>6</td>
<td>Binomial Distribution</td>
<td>Student choice of interesting topics</td>
<td>N/Aa</td>
</tr>
<tr>
<td>7</td>
<td>Poisson Distribution</td>
<td>Leisure Activities &amp; Hobbies - Conspiracy Theories</td>
<td>73.5%</td>
</tr>
<tr>
<td>8</td>
<td>Uniform Distribution</td>
<td>Technology - TV Binge Watching / Screen Time</td>
<td>86.7%</td>
</tr>
<tr>
<td>9</td>
<td>Normal Distribution</td>
<td>Sports &amp; Activities - Exercise / Working out</td>
<td>89.2%</td>
</tr>
<tr>
<td>10</td>
<td>Exponential Distribution</td>
<td>Sports &amp; Activities - Spectator Sports</td>
<td>69.9%</td>
</tr>
</tbody>
</table>

*Note.* Interest Survey Percentage represents the percent of students who rated the relevance of the context as a 3 or higher out of 5.

*a Interest Survey Percentage is not applicable because students researched personal items of interest in the activity, and the context was not uniform for all students.*
**Intervention Fidelity.** Exit slips were viewed only by the course instructors and the researcher who was also one of the course instructors. The relevance rating for each student was recorded on a spreadsheet with names replaced by unique identifiers prior to disclosure to the rest of the research team. Some students were not writing the appropriate context in the blank space on their exit slip after the first three activities, but were instead filling in the blank with the statistical content of the lesson. This led the instructor-researcher to believe that some students were misunderstanding the directions and rating the content of the lesson rather than the context for the activity. As a result, before the fourth exit slip was filled out, the instructor-researcher verbally clarified the difference between content and context to all sections and asked students to only rate the context of the activity. After this clarification, no further misrepresentations were noticed on the exit slips.

**Results.** Exit slips were collected after each activity and the average relevance rating was calculated. Except for the first three activities, the treatment group consistently had higher average relevance ratings than the control group (see Table 4.2). When taken altogether, the average relevance rating for the last seven activities of the treatment group was 3.98, while the average relevance rating for the last seven activities of the control group was 2.74.
Table 4.2

Relevance Ratings of Context for Intervention Activities

<table>
<thead>
<tr>
<th>Lesson Content</th>
<th>Treatment</th>
<th>Control</th>
<th>Treatment</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity 1 - Basic Probability and Law of Large Numbers</td>
<td>Jimmy Fallon’s Egg Roulette Game</td>
<td>3.86 ($n=33$)</td>
<td>Dice Rolling Game</td>
<td>4.27 ($n=22$)</td>
</tr>
<tr>
<td>Activity 2 - Addition Law of Probability</td>
<td>Travel Destinations</td>
<td>4.15 ($n=34$)</td>
<td>Junk Food</td>
<td>4.25 ($n=20$)</td>
</tr>
<tr>
<td>Activity 3 - Tree Diagrams and Multi-stage Probabilities</td>
<td>Birth Control</td>
<td>3.71 ($n=26$)</td>
<td>Widgets</td>
<td>3.90 ($n=20$)</td>
</tr>
<tr>
<td>Activity 4 - Conditional Probability</td>
<td>Pets and Dating Desirability</td>
<td>3.93 ($n=30$)</td>
<td>Worms</td>
<td>2.04 ($n=23$)</td>
</tr>
<tr>
<td>Activity 5 - Discrete Variables and Expected Value</td>
<td>Batch Testing and COVID</td>
<td>3.87 ($n=30$)</td>
<td>Rocks and Mining</td>
<td>2.00 ($n=21$)</td>
</tr>
<tr>
<td>Activity 6 - Binomial Distribution</td>
<td>2 Truths and a Lie Game</td>
<td>4.28 ($n=18$)</td>
<td>Rock, Paper, Scissors</td>
<td>3.56 ($n=16$)</td>
</tr>
<tr>
<td>Activity 7 - Poisson Distribution</td>
<td>Conspiracy Theories</td>
<td>3.97 ($n=35$)</td>
<td>Pipeline Leaks</td>
<td>2.33 ($n=21$)</td>
</tr>
<tr>
<td>Activity 8 - Uniform Distribution</td>
<td>The Price is Right Games</td>
<td>3.67 ($n=26$)</td>
<td>Number Games</td>
<td>3.05 ($n=20$)</td>
</tr>
<tr>
<td>Activity 9 - Normal Distribution</td>
<td>Weightlifting</td>
<td>4.13 ($n=31$)</td>
<td>Baby Birth Weight and Dice</td>
<td>3.30 ($n=20$)</td>
</tr>
<tr>
<td>Activity 10 - Exponential Distribution</td>
<td>NCAAA March Madness</td>
<td>4.08 ($n=24$)</td>
<td>WWII Planes</td>
<td>3.26 ($n=19$)</td>
</tr>
</tbody>
</table>

Note. Clarification was made to the students about the difference between content and context prior to the 4th exit slip.
Additionally, exit slips were used as a way to track participant attendance throughout the intervention. Of the 43 students initially in the treatment group, two did not participate in any of the intervention activities. A total of 30 students within the treatment group participated in at least six, or the majority, of the activities. Of the 29 students initially in the control group, two did not participate in any intervention activities, and 23 participated in a minimum of six activities (see Appendix E.2). Participants were deemed eligible for quantitative data analysis if they attended at least one activity during the intervention. Therefore, 41 students in the treatment group, and 27 students in the control group contributed to the quantitative analysis of comprehension and attitude. In order to be eligible for the qualitative interviews and focus groups, participants needed to attend more than half of the activities in the intervention, although priority was given to students with higher attendance rates where possible.

An average relevance score was calculated for the last seven activities for each of the 68 students attending at least one intervention activity (see Table 4.3). Only one student from the treatment group had an average relevance score below 3 (2.4%), three had an average relevance score of exactly 3 (7.3%), and the remaining 37 students had an average relevance score higher than 3 (90.2%). In the control group, 18 out of 27 students had an average relevance score below 3 (66.7%), two students had an average relevance score of exactly 3 (7.4%), and the remaining seven students had an average relevance score higher than 3 (25.9%). That is, most students in the treatment group found the context in the activities personally relevant whereas the majority of students in the control group did not.
Table 4.3

*Individual Average Relevance Ratings for the Last Seven Activities*

<table>
<thead>
<tr>
<th>Average Relevance over Last Seven Activities</th>
<th>Number of Treatment Group Students</th>
<th>Number of Control Group Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2.999</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>Exactly 3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>3.001-5</td>
<td>37</td>
<td>7</td>
</tr>
</tbody>
</table>

*Comprehension (Research Question 1)*

**Data Collection.** In an attempt to determine the extent to which using relevant context affected student comprehension of probability concepts, students were given a 10-question pretest during the week prior to the beginning of the probability unit and a two-part posttest during the week after the conclusion of the unit. The posttest contained the same 10 pretest questions and additional questions to ensure all key topics covered in the unit were assessed. Approximately half of the questions on both the pretest and posttest measured statistical literacy and the other half measured statistical reasoning. For the purposes of this research, comprehension of statistical concepts is measured by both statistical literacy and reasoning.

Upon completion of the unit, student comprehension was measured using two different strategies. First, the same 10 questions from the pretest were embedded within the posttest, and the change in score on these questions was calculated for each participant. Second, each student received a holistic score out of 100 on their posttest, and the average student score was compared between the treatment and control groups.

**Intervention Fidelity.** Some participants were eliminated from the comprehension portion of the study for not completing both the pretest and posttest. The
treatment group sample size, which was initially composed of 41 participants who attended at least one activity, lowered to 38 participants after three of them failed to take the posttest. The control group sample size decreased from 27 participants to 25 participants for the same reason. Additionally, it was only possible to obtain the overall posttest score for one of the treatment group students, not the granular test data needed to determine change from pretest to posttest score. Therefore, for the gain in pretest score calculations, the treatment group sample size was 37.

**Baseline Data.** Prior to the analysis to answer the research question, the pretest scores were first analyzed to determine if there was any significant difference in background knowledge between treatment and control groups. The 10-question pretest consisted of six multiple choice questions and four fill-in-the-blank questions. Responses for each multiple choice question was given a value of 1 if answered correctly, or 0 if answered incorrectly. All of the fill-in-the-blank questions were also scored with this 1-point scale. However, three of the four fill-in-the-blank questions allowed for a half-point if a student was able to demonstrate a partial understanding of the correct computational process. For the purposes of analyzing pretest scores, the potential participant values ranged from a minimum of 0 (all questions answered incorrectly) to a maximum of 10 (all questions answered correctly).

The mean scores on the 10-question pretest for the treatment group and control group were similar. The mean score for the treatment group was 2.61 with a standard deviation of 1.31, while the mean score for the control group was 2.38 with a standard deviation of 1.42. Both groups had students who scored 0 points. The highest score for the treatment group was 5.5, but the control group had a high score of 5 (see Table 4.4).
Further analysis of potential differences in background knowledge was conducted using the boxplots shown in Figure 4.1. Both groups had similar medians, but the treatment group had a larger spread, as seen by the difference in interquartile range.

**Figure 4.1**

_Distribution of Pretest Scores_

Next, a two-sample $t$-test was used to determine if there was a significant difference in background knowledge between the treatment and control groups based on the 10-question pretest scores. Conditions of normality, equal variance, and independence were met, and a $p$-value of 0.5195 was obtained, which is greater than the preset significance level of 0.05. Therefore, it follows that the amount of background
knowledge of the students in the treatment and control groups were not significantly different from each other based on pretest scores. Because no significant difference was found between the groups, it can be assumed that any differences between treatment and control groups, based on the posttest, would be attributed to the intervention or to external factors rather than to the prior knowledge brought into the course.

**Results.** Two separate measures were considered in an attempt to measure changes in student comprehension following the intervention. First was the change in student scores for the 10 pretest questions which were also embedded in the posttest. Second, overall posttest scores were considered. Normality conditions were met for both measures, as were equal variances and independence. Therefore, a two-sample *t*-test was used to analyze both measures. Results are shown in Table 4.5.

**Table 4.5**

*Post-Intervention t-test Results for Comprehension Measures*

<table>
<thead>
<tr>
<th></th>
<th>Treatment Group</th>
<th></th>
<th>Control Group</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td><em>p</em>-Value</td>
</tr>
<tr>
<td>Change in Pretest Scores</td>
<td>2.92 (n=37)</td>
<td>2.27</td>
<td>2.88 (n=25)</td>
<td>1.86</td>
<td>0.9436</td>
</tr>
<tr>
<td>Overall Posttest Scores</td>
<td>64.37 (n=38)</td>
<td>19.88</td>
<td>60.12 (n=25)</td>
<td>19.93</td>
<td>0.4104</td>
</tr>
</tbody>
</table>

*Note.* Change in pretest scores was calculated as (posttest score - pretest score) with possible values ranging between -10 and 10. Possible overall posttest score values range from 0 to 100.

**Change in Pretest Scores.** The same 10 questions that were on the pretest were embedded into the posttest, and the same 10-point scale was applied when analyzing the change in score from pretest to posttest. The change in pretest score was calculated by
taking posttest score minus pretest score. Therefore, potential participant values for the change in pretest score ranged between -10 and 10.

Change in score was first compared using descriptive statistics and a box plot (see Table 4.6 and Figure 4.2). The mean change in scores on the 10-question pretest for the treatment and control groups were similar. Both groups answered approximately three more questions correctly, on average, when taking the posttest compared to the pretest. The mean change in score for the control group was 2.88 with a standard deviation of 1.86, while the mean change in score for the treatment group was just slightly higher at 2.92 with a standard deviation of 2.27. Both groups had students who scored worse on the posttest than on the pretest, as indicated by the negative minimum values in Table 4.6.

Gains on the posttest were defined as any change higher than 0. A similar percentage of both groups showed gains on the posttest. 32 out of 37 treatment group participants (86.5%) and 22 out of 25 control group participants (88%) earned higher scores on the posttest than on the pretest. The highest individual gain for the treatment group was 8, while the control group had an individual gain as high as 7 (see Table 4.6).

Table 4.6

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Max</th>
<th>% with Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>37</td>
<td>2.92</td>
<td>2.27</td>
<td>-1.5</td>
<td>8</td>
<td>86.5%</td>
</tr>
<tr>
<td>Control</td>
<td>25</td>
<td>2.88</td>
<td>1.86</td>
<td>-0.5</td>
<td>7</td>
<td>88%</td>
</tr>
</tbody>
</table>

Note. % with Gains is the percentage of students with a change in score greater than 0.

Boxplots were used to compare the change in scores from pretest to posttest (see Figure 4.2). Both groups had a median change of 3. Although there was slightly more
spread in the treatment group, both boxplots indicate that the data is relatively symmetric and spread similarly.

**Figure 4.2**

*Boxplots of Posttest and Change from Pretest*

To help answer the first research question, a two-sample *t*-test was conducted to determine whether the average change in pretest score was significantly different between the treatment and control group. All three assumptions of independence, normality, and equal variances held for the change in pretest scores, and therefore the conditions necessary for a two-sample *t*-test were met. Upon completion of the *t*-test, a *p*-value of 0.9436 was obtained, which is far greater than the significance level of 0.05 (see Table 4.5). It follows that the change in scores from pretest to posttest of the treatment and control groups were not significantly different from each other. In other words, the use of relevant context did not appear to create a significant difference in student comprehension, as assessed by the change in pretest scores.

**Overall Posttest Scores.** As a final analysis on comprehension, data on the overall posttest scores were compared between the treatment and the control group. From descriptive statistics computations, the mean posttest scores for the treatment group was
slightly higher than the control group. The mean score for the treatment group was 64.37 with a standard deviation of 19.88, while the mean score for the control group was lower at 60.12 with a standard deviation of 19.93. The treatment group had a larger spread in overall scores than the control group. Scores from the treatment group ranged in value between 12 and 100, while the control group saw scores between 25 and 90 (see Table 4.7).

**Table 4.7**

*Descriptive Statistics for Posttest Scores*

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>38</td>
<td>64.37</td>
<td>19.88</td>
<td>12</td>
<td>100</td>
</tr>
<tr>
<td>Control</td>
<td>25</td>
<td>60.12</td>
<td>19.93</td>
<td>25</td>
<td>90</td>
</tr>
</tbody>
</table>

Because all necessary conditions were met, a two sample *t*-test was used to determine whether there was a difference in average posttest scores between the treatment and control groups. A *p*-value of 0.4112 was obtained, which is greater than the preset significance level of 0.05 (see Table 4.5). Posttest averages between the treatment and control groups were found to not be significantly different. Thus, the use of relevant context did not appear to have created a significant difference in comprehension.

**Results Based on Attendance.** Due to the fact that some students were absent on most days during the probability unit and not exposed to many of the intervention activities, an additional analysis was conducted for the set of students who participated in at least six of the 10 activities. *T*-tests were run on the average change in pretest scores and on the average overall posttest scores based on the minimum number of activities attended (see Table 4.8). Every calculated *p*-value was greater than the significance level
of 0.05. Therefore, it was concluded that there were no significant differences in the changes on the pretest score or on the overall posttest scores when comparing the treatment to the control group.

**Table 4.8**

*P-values Based on Number of Activities Attended*

<table>
<thead>
<tr>
<th>Minimum Number of Activities Attended</th>
<th>Change in Pretest Score <em>p</em>-value</th>
<th>Posttest Score <em>p</em>-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.8883 (49 df, <em>t</em> = 0.14)</td>
<td>0.3428 (50 df, <em>t</em> = -0.98)</td>
</tr>
<tr>
<td>7</td>
<td>0.9748 (44 df, <em>t</em> = -0.03)</td>
<td>0.4838 (45 df, <em>t</em> = -0.71)</td>
</tr>
<tr>
<td>8</td>
<td>0.9923 (32 df, <em>t</em> = 0.01)</td>
<td>0.7450 (33 df, <em>t</em> = -0.33)</td>
</tr>
<tr>
<td>9</td>
<td>0.6616 (24 df, <em>t</em> = -0.44)</td>
<td>0.3277 (24 df, <em>t</em> = -1.00)</td>
</tr>
<tr>
<td>10</td>
<td>0.9261 (13 df, <em>t</em> = 0.09)</td>
<td>0.4702 (13 df, <em>t</em> = -0.74)</td>
</tr>
</tbody>
</table>

**Attitude (Research Question 2)**

**Data Collection.** To determine whether the intervention led to a significant change in student attitude toward statistics, participants were first given the pre-version of the SATS-36 before the beginning of the intervention. Each student received six attitude subscores, which were obtained by averaging student Likert ratings on survey questions corresponding to each of the six attitude components. Scores were given on a scale from 1 to 7, with 1 representing a very negative attitude, 4 indicating a neutral attitude, and 7 indicating a very positive attitude. Upon completion of the intervention but before the unit exam, students were given the post-version of the SATS-36, and change in each of the six attitude components was calculated for each student.

**Intervention Fidelity.** Of the 41 students eligible for participation in the treatment group, a total of 38 completed both versions of the SATS-36, and were thus
included in the data analysis. In the control group, there were 25 students out of a possible 27 who completed both versions of the SATS-36. Attendance rates of students varied widely during the intervention period, with students attending anywhere from none to all of the 10 activities. Although lower attendance means less exposure to the relevant context intervention, the attitude analysis does not take this factor into account.

**Baseline Data.** After the administration of the pre-version of the SATS-36, summary statistics for the treatment and control group were calculated, and analysis was conducted to determine whether a statistically significant difference existed between the two groups. Not all components met the normality requirements needed for a two-sample t-test, so all six attitude components were analyzed using the non-parametric Mann-Whitney U test instead. P-values were calculated to compare the treatment and control group distributions for each of the six attitude components (see Table 4.9).

**Table 4.9**

<table>
<thead>
<tr>
<th>Attitude Component</th>
<th>Treatment Group (n=38)</th>
<th>Control Group (n=25)</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Affect</td>
<td>4.43</td>
<td>1.38</td>
<td>4.27</td>
</tr>
<tr>
<td>Cognitive Competence</td>
<td>5.04</td>
<td>1.08</td>
<td>5.13</td>
</tr>
<tr>
<td>Value</td>
<td>5.18</td>
<td>1.04</td>
<td>5.16</td>
</tr>
<tr>
<td>Difficulty</td>
<td>3.34</td>
<td>0.95</td>
<td>3.39</td>
</tr>
<tr>
<td>Interest</td>
<td>5.19</td>
<td>1.40</td>
<td>4.98</td>
</tr>
<tr>
<td>Effort</td>
<td>6.47</td>
<td>0.52</td>
<td>6.44</td>
</tr>
</tbody>
</table>

*Note.* Rated on a 7-point Likert Scale (1=Strongly Negative, 7=Strongly Positive).
Averages across five of the six attitude components for both groups were greater than 4.0, implying that students in both groups entered the course with at least a slightly positive attitude toward statistics. In both groups, the component of Effort scored the highest, at 6.47 and 6.44 for the treatment and control groups respectively. This suggests that students in both groups entered the course with strong intentions to put their best effort into their work. The only component of the six that did not score positively was that of Difficulty, which scored at 3.34 and 3.39 for the treatment and control groups respectively. This suggests that students anticipated having at least some challenge with the material covered in the course. $P$-values across all six group comparisons were well over the 0.05 significance threshold used. As a result, no strong evidence was found to suggest any initial differences in attitude that existed between the treatment and control group. Further analysis of potential differences in initial student attitude was conducted using the boxplots shown in Figure 4.3.

**Figure 4.3**

*Distribution of Pre-Attitude Scores for Students in Treatment and Control Group*
Similar to the averages compared above, median responses between the treatment and control group did not differ greatly in any of the six attitude components. A majority of students in both the treatment and control group entered the course with a positive Affect and Interest toward statistics, which was contrasted with results found post-intervention. In terms of variability of data, the treatment group, which was composed of two separate sections of the course, does exhibit greater variability than the control group, with substantially larger interquartile range for the components of Affect, Cognitive Competence, Value, and Interest. A small number of outliers are visible in the boxplots for each of the groups.

**Results.** Upon conclusion of the six-week relevant context intervention, but before taking the unit exam, students were given the post-version of the *SATS-36*. Summary statistics were calculated as with pre-intervention attitude, and boxplots were constructed. However, *p*-values were not calculated, as the post-analysis was performed on individual student change in attitude rather than their post-intervention attitude (see Table 4.10).

Average student score in post-intervention attitude decreased across 11 of the 12 calculated means. The only exception to this was the Difficulty rating for the control group, which remained identical at 3.39. While none of the categories decreased enough to switch from a positive to a negative attitude, the change in Affect toward Statistics pushed both treatment and control groups from a slightly positive rating (4.43 and 4.27, respectively) to a neutral rating (3.98 and 4.04, respectively). The distribution of individual student post-intervention attitude is visible in Figure 4.4.
Table 4.10

Post-Intervention Attitude Averages

<table>
<thead>
<tr>
<th>Attitude Component</th>
<th>Treatment Group (n=38)</th>
<th>Control Group (n=25)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Affect</td>
<td>3.98</td>
<td>1.46</td>
</tr>
<tr>
<td>Cognitive Competence</td>
<td>4.47</td>
<td>1.38</td>
</tr>
<tr>
<td>Value</td>
<td>4.78</td>
<td>1.32</td>
</tr>
<tr>
<td>Difficulty</td>
<td>3.22</td>
<td>1.10</td>
</tr>
<tr>
<td>Interest</td>
<td>4.70</td>
<td>1.69</td>
</tr>
<tr>
<td>Effort</td>
<td>5.89</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Note. Rated on a 7-point Likert Scale (1=Strongly Negative, 7=Strongly Positive).

Figure 4.4

Distribution of Post-Attitude Scores for Students in Treatment and Control Group
The distribution of post-intervention attitude shows a greater deal of variability than the pre-intervention attitude data. Of the 12 distributions, 11 showed a larger range than their pre-intervention attitude counterparts, and 10 exhibited a greater interquartile range. This increase in variability could suggest that the effects of the intervention on student attitude varied greatly across different participants.

Next, the change in attitude was calculated for the six attitude components of each student by taking their respective post-attitude score and subtracting the corresponding pre-attitude score. To determine whether a significant difference in attitude change occurred between the treatment and control groups, summary statistics were calculated, along with $p$-values from the comparative Mann-Whitney U tests for each of the six attitude components (see Table 4.11).

**Table 4.11**

*Change in Attitude Averages*

<table>
<thead>
<tr>
<th>Attitude Component</th>
<th>Treatment Group ($n=38$)</th>
<th>Control Group ($n=25$)</th>
<th>$p$-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affect</td>
<td>-0.45 (SD=1.25)</td>
<td>-0.23 (SD=1.17)</td>
<td>0.5645</td>
</tr>
<tr>
<td>Cognitive Competence</td>
<td>-0.57 (SD=0.93)</td>
<td>-0.35 (SD=1.02)</td>
<td>0.4313</td>
</tr>
<tr>
<td>Value</td>
<td>-0.40 (SD=1.05)</td>
<td>-0.22 (SD=0.74)</td>
<td>0.5837</td>
</tr>
<tr>
<td>Difficulty</td>
<td>-0.12 (SD=0.90)</td>
<td>0.00 (SD=0.66)</td>
<td>0.7254</td>
</tr>
<tr>
<td>Interest</td>
<td>-0.49 (SD=1.45)</td>
<td>-0.70 (SD=0.83)</td>
<td>0.4452</td>
</tr>
<tr>
<td>Effort</td>
<td>-0.58 (SD=0.92)</td>
<td>-0.35 (SD=0.84)</td>
<td>0.2647</td>
</tr>
</tbody>
</table>

*Note.* Change in each attitude component was calculated by (postscore - prescore).
As seen in the post-attitude description, averages for 11 of the 12 components are negative, implying that student attitude actually decreased over the course of the intervention, for both students in the treatment and control group. However, these decreases did not differ significantly between the treatment and control group, as evidenced in the \( p \)-values that are all well above the 0.05 significance threshold. This suggests that the decrease in attitude was likely caused by another factor, such as the difficulty of the probability concepts covered during the unit. There does not appear to be a statistically significant difference in student attitude created by the relevant context intervention. Further analysis of the distribution is possible using the boxplots shown below in Figure 4.5.

**Figure 4.5**

*Distribution of Attitude Change for Students in Treatment and Control Group*

In agreement with the negative averages shown above, the median change in attitude is at or below zero for all six attitude components in both the treatment and
control group. In other words, over 50% of students in both groups experienced a
decrease in each component of their attitude toward statistics. In analyzing the variability
of the data, the components of Affect and Interest demonstrated substantially larger range
and interquartile range for the treatment group, in comparison to the control group.
Because the treatment group was composed of two sections of the course taught by
different instructors, further investigation was conducted to see if any differences existed
in attitude between the three different sections used in the study.

Results Separated by Instructor. A brief analysis of attitude data separated by
instructor was conducted, to determine whether any differences existed between
individual sections used for the study. Notable differences were found within the attitude
components of Affect and Interest, which have been included in more detail below. The
remaining four attitude components were not found to be noteworthy and have not been
included in this additional analysis. Figure 4.6 reveals differences in Affect and Interest
respectively between the three sections under study.

Figure 4.6

Change in Affect and Interest Separated Between Class Sections
In this figure, the single control group section remains unchanged from earlier graphs. However, results from the treatment group have been split into two separate sections, Treatment Group 1 and Treatment Group 2. Treatment Group 1 was taught by a member of the research team with 17 years of experience. Treatment Group 2 and the control group were taught by two newer staff members with three years of experience each. The median change in attitude for Treatment Group 1 was greater than that of Treatment Group 2 and the control group for both Affect and Interest, with a higher percentage of students experiencing a positive change in these attitude components.

According to the primary creator of the SATS-36, a change in an attitude component by 0.5 or more can be considered important (Schau & Emmioglu, 2012). The percentage of students in each section of the course who met this threshold, both in a positive and negative direction, was calculated, and is shown in Tables 4.12 and 4.13.

**Table 4.12**

*Percent of Students with Significant Change in Affect*

<table>
<thead>
<tr>
<th></th>
<th>Control Group (n=25)</th>
<th>Treatment Group 1 (n=20)</th>
<th>Treatment Group 2 (n=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase of 0.5 or More</td>
<td>20.0%</td>
<td>35.0%</td>
<td>16.7%</td>
</tr>
<tr>
<td>Decrease of 0.5 or More</td>
<td>52.0%</td>
<td>40.0%</td>
<td>55.6%</td>
</tr>
</tbody>
</table>

**Table 4.13**

*Percent of Students with Significant Change in Interest*

<table>
<thead>
<tr>
<th></th>
<th>Control Group (n=25)</th>
<th>Treatment Group 1 (n=20)</th>
<th>Treatment Group 2 (n=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase of 0.5 or More</td>
<td>8.0%</td>
<td>25.0%</td>
<td>11.1%</td>
</tr>
<tr>
<td>Decrease of 0.5 or More</td>
<td>64.0%</td>
<td>30.0%</td>
<td>66.7%</td>
</tr>
</tbody>
</table>
Substantially more students emerged from Treatment Group 1 (35.0%) with a large positive change in Affect when compared to Treatment Group 2 (16.7%) and the control group (20.0%). In contrast, fewer students from Treatment Group 1 exhibited a substantial negative change in Affect (40.0% versus 55.6% and 52.0% respectively). A similar difference between sections is visible in student Interest toward statistics. In Treatment Group 1, 25.0% of students improved substantially in their Interest, compared to 11.1% in Treatment Group 2, and only 8.0% in the control group. While nearly two thirds of students in Treatment Group 2 (66.7%) and Control (64.0%) experienced a large decrease in their Interest toward statistics, only 30.0% of students in Treatment Group 1 displayed a similar decrease. This difference in student Affect and Interest is consistent with earlier findings by researchers such as Hattie (2015) and Di Martino & Zan (2010) suggesting the importance of instructor in student outlook toward a course.

Qualitative Findings

Factors Affecting Comprehension and Attitude (Research Question 3)

Data Collection. Deductive category coding was used to answer the third research question. This approach began the coding process with an initial set of codes. The findings are based on the analysis of one-on-one interviews and focus groups of students in the treatment and control groups. Twelve transcripts were analyzed to uncover themes to better understand the results from the first two research questions. A total of 16 students participated in the interviews and focus groups. There were 10 one-on-one interviews with three students from Treatment Group 1, three students from Treatment Group 2, and four students from the control group. Two focus groups were conducted. The focus group from Treatment Group 1 included three students and the focus group for
the control group included two students. Schedule conflicts prevented a focus group for Treatment Group 2 from occurring. Students with 80% or greater attendance were listed as a priority for participation in the qualitative portion of the study. Additionally, students with an attendance rate of 50% or less were removed from the list of potential candidates.

**Interview and Focus Groups Protocol.** The interviews and focus groups were conducted via Zoom, a cloud-based video conferencing platform. A semi-structured format was used which allowed for the same questions during the interview and focus groups with open ended responses and additional follow-up questions included, as necessary. The Zoom sessions were recorded then transcribed. Participant anonymity was protected by having records stored on a secure password-protected computer and by having transcripts stripped of any identifiable information from participants. A unique identifier only known by the researcher conducting the interviews and focus groups was used instead of participant names on the transcripts.

**Analysis.** A deductive approach was used for the analysis as the researchers started with categorical codes based on the first two research questions and the theoretical framework. The in vivo coding method defined the initial coding list from the categorical codes. The in vivo coding method uses the actual words and phrases from the interview participants (Yin, 2016).

The transcripts were coded manually based on this initial code list. An iterative comparative analysis was used to revise and update the coding list. Transcripts were then uploaded into the computer software, NVivo, for additional analysis and for a check of reliability. Additionally, a separate reliability check of codes was made by the other researchers to verify that no additional codes were needed.
The primary coding researcher supplied the other researchers with anonymized transcripts. The other researchers then coded the transcripts using the codebook, and results were compared to ensure fidelity and consistency in coding. Based on the collective coding, a final code list (see Appendix E.3) was determined in which all the transcripts were coded again.

Using the NVivo software, the reference count was used as a tool to verify the frequency of the codes to determine the emerging themes. The sections that follow indicate the themes from each of the four categories (two in Figure 4.7 and two in Figure 4.8): Factors that Affected Comprehension, The Effect of Context on Comprehension, Factors that Affected Attitude, and The Effect of Context on Attitude.
Figure 4.7

Codes to Theme Illustration for Factors Affecting Comprehension

Categorical Code

- The flipped classroom setting was not useful to me in my setup
- The flipped classroom is helpful in that I can actually come in with questions

In Vivo Code

- Groups are a struggle for me
- Group activities are discouraging
- Like doing things in groups to reinforce my comprehension

Theme

Elements of Instructional Process

- The videos break it down
- Not one that can learn math from videos
- Video lectures are good because I can go back and rewatch it

Class Structure

- It is a lot of word problems
- It’s too much
- A lot more being added on
- Symbols being reused in different ways

Amount of Material

- The instructor explains the material
- Watching what my instructor did as they went through the problem
- Instructor does a good job answering questions

Instructor

Outside Circumstances

- School is dependent on my kids’ schedule
- Home stuff overwhelm school time
- A lot of things get in the way of me trying to get the time

Schedule

- With my work schedule, it felt like school was put on the backburner
- Work overnights so do assignments at night
- Couldn’t because of work schedule
Factors that Affected Comprehension. Questions asked during the interviews focused on two categories: the instructional process and any outside circumstances that students felt might have affected their comprehension. Figure 4.7 shows the themes formed from both of the categories along with examples of the in vivo coding that were representative of these themes. Participants were also asked specifically how the context used during the in-class activities affected their comprehension. Four themes emerged about factors affecting comprehension, three associated with the elements of the instructional process and one from the outside circumstances:

Category: Elements of Instructional Process. This category covers aspects of the instruction of the unit including video lectures, in-class activities, peer support, classroom logistics, and the instructor. There were 18 codes that were assigned to this category. The themes that emerged were mentioned by at least 70% of those interviewed.

Theme 1: Classroom Structure. The majority of codes associated with this theme include flipped classroom, video lectures, and groups for in-class activities. These codes were referenced 27 times. Additional codes included practice programs, lecture notes, and student instructor. The flipped classroom setting was mentioned in the majority of those interviews as affecting how students learned the material. “My honest opinion about the flipped class setting is that I would really prefer to go over everything in class rather than have a flipped classroom setting,” said one student interviewed from the treatment group. Another student from the treatment group confirmed this factor by saying “I am able to
comprehend the material eventually. It's just I wouldn't say I have immediate comprehension due to the flipped classroom setting.” They went on to share that the group work helped them “form stronger bonds with people in the class that [they] would not have known otherwise.” This allowed them to ask questions which had “definitely been crucial and increased [their] understanding.”

Other students referenced the classroom structure and flipped classroom by referring to the video lectures as part of the instructional process. There were differing thoughts. A student from the treatment group stated, “I learned best through trial and error and asking the professor questions so, generally speaking, video lectures just don’t work for me.” A student from the control group did find the video lectures useful but disliked the classroom activities that followed. “Okay, video lectures are really good for me because I keep on rewatching them. Oh, I would say [the classroom activities] are at the bottom of my list because the days that I want more one-on-one are the days that we do that in classroom activity. I sometimes think the classroom activities are a waste of time.”

Theme 2: Amount of Material. This theme encompasses the responses associated with the probability content and topics. There were 17 references coded as “a lot of material.” Additional codes included the difficulty of the material, the formulas, and the formatting of questions. This quote from a student in the treatment group summarizes the
responses: “The topic started to pick up and we were taking the same amount of time, as you were on the previous skills, but I felt like the topics were harder for me to understand. I had to spend more time so that is when my comprehension started to go down. I just felt like there was not enough actual time in class to have that information digest.”

Theme 3: Instructor. The instructor provided the in-class instruction of the curriculum. More than 75% of interviewees shared responses on how the instructor affected their comprehension of the probability material. Students from both the treatment group and the control group shared how the instructor and their actions affected how they comprehended the material. “I have had professors who didn't know what they were talking about, but my instructor explained it really well,” said a student in the treatment group. Another student from the treatment group shared, “I think back to watching what my instructor did as [they] went through the problem and I try to remember each step.” A student from the control group also mentioned the effects of the instructor by stating “It helps me a lot when [the instructor] goes over stuff we should have learned and reinforces them through activities.”

Category: Outside Circumstances. This category examined the circumstances outside of the instructional process that could have affected the comprehension of the probability unit material. There were notably fewer responses to this inquiry with less than 40% of those interviewed providing coded responses. The codes that emerged in this category include learning disabilities,
prior knowledge from having taken the class previously, family, sickness, and time commitments to work.

Theme 1: Schedule. Schedule was the emerging theme from outside circumstances that affected material comprehension. This incorporates the codes of time, work, and family. As such, schedule is defined as time being a factor because of family, work, or life situations. A student from the control group shared, “It is harder for me timewise to do the extra stuff, so I value the in-class time. That is 100% set aside with no distractions.” This same feeling about time emerged in the treatment group as highlighted by this quote, “Five kids and I don't have any help because my husband goes out of town for work. I have no family here so there are a lot of things that get in the way of me getting the time.”

The Effect of Context on Comprehension. Participants were asked directly about how the context used during the in-class activities affected how they comprehended the concepts. This question was a direct link to research question one and to the theoretical framework on relevant context affecting comprehension. Codes under this category included the material being “relatable”, being “aligned with career or major”, and the context being “engaging”. Another code under this category included the context “not making a difference” to comprehension. A majority of responses indicated that the context was interesting to students. The notion of ”interesting” meant that the context was fun, cool, hands-on, or memorable. While participants commented on the context being interesting, there were differing viewpoints on if the context was helpful to the comprehension of the material.
The notion that the relevant contexts were “interesting” emerged during the interviews and focus groups. Participants were asked to define what “interesting” meant in regard to their learning. These quotes, compiled from treatment group and control group interviews, capture the essence of the findings. “I think [the contexts] kept my focus. I would not really call it interesting enough to make me want to love statistics, but it definitely kept my attention for that class,” said a student from the treatment group. A student from the control group shared, “Personally, the topics that I found more interesting I feel make me kind of want to do it more and I may pay more attention. If it's interesting and it holds my focus, I hear more, I comprehend more, and I catch more. Instead of just spacing out and just hearing the words, I actually listen to the words.”

There were varying opinions on the usefulness and helpfulness of the classroom activities that used relevant context. There were 20 references (14 from the treatment group and six from the control group) to the context being helpful, engaging, and improving comprehension. There were 28 references (12 from the treatment group and 16 from the control group) to context being not very helpful, confusing, silly, a waste of time, fleeting, or not making a difference to the comprehension. The activities used in the treatment group had contexts tailored to student interests, as determined by survey responses. The contexts used for the control group were not tied to student interests.

Students in the treatment group noted that the contexts were helpful although they still required more instruction: “The real-world examples almost confused me because I need to understand the base level concept. The context is really helpful when you know what all the different pieces mean and then see if you can apply it, but if you don't have the foundation, the real world is confusing,” shared by one student. Another student
stated that they “appreciate that the context puts the information into perspective that the student can understand. I do not think the specific context used made much of a difference. There being context really demonstrated how it actually works as opposed to just numbers. Statistics is a math but it's a very social math, so the context is very helpful because it allows you to understand what the purpose of it is.”

Students from the control group had different thoughts on the in-class activities and their value in the instructional process and comprehension of the probability material. The quotes from one student in the control group highlighted the difference they experienced. Initially, the student commented on how the context being used “felt comical because I would be feeling how earthworms relate to your everyday life or how rocks really relate. They do not. The main difficulty was nobody really knew what they were doing so it's kind of hard to concentrate on something when you're just staring at your paper for like 20-30 minutes.” The student then shared the difference when their professor provided some real-world context to one of the problems. “He would use examples of restaurants, food change, drink choices, or just things that were more relatable whenever he would write on the board. I thought that was easier to follow than if he was doing an activity about rocks or something like that.”

The use of the relevant context being coded as improving comprehension was referenced only four times. One student in the treatment group shared that relevant context “definitely improved [their] comprehension over how [they] would have done in a normal 100% old school math style textbook.” A student in the control group correlated how they connect context to comprehension. “When I am struggling with a problem and trying to remember how it all works, I try to remember the context. I try to remember if
this was the rocks on a cart problem or this was the candy bars on the factory line or this was the baby weight. I try to recall the situation I was in when I was learning it to remind myself how it all works. Having those setups makes it better for me to be able to recall that information.”

Factors That Affected Attitude. Another round of questions asked during the interviews focused on the factors that affected attitude. Figure 4.8 shows the themes formed from each of the categories of the instructional process and any outside circumstances which may have affected attitude. Participants were also asked specifically about how the context used during the in-class activities affected their attitude, if at all. Two themes emerged from the questions associated with the elements of the instructional process that affected attitude. There were more references to outside circumstances affecting attitude than any other category.

Category: Elements of Instructional Process. This category covers aspects of instruction, including the feedback and attitude of the instructor in addition to in-person classes versus online instruction. There were 17 codes that were assigned to this category. The emerging themes were mentioned by at least 60% of those interviewed.
Figure 4.8

*Codes to Theme Illustration for Factors Affecting Attitude*

**Categorical Code**
- Having interactions and being able to talk to the teacher makes class more enjoyable
- The formula sheet helped take a little stress off

**In Vivo Code**
- Struggled a lot with online classes
- The classes are good
- More negative if had to take the class strictly online
- Not positive because class is hard
- No idea what we are talking about so attitude changed
- It started high and then the quizzes got harder so it went down

**Theme**
- Class Structure
- Instructor
- Exam Performance
- Previous Experiences

**Elements of Instructional Process**
- Probably my teacher and the way the lectures are conducted
- The instructor is always on topic
- Getting immediate feedback helps a lot
- Getting reassurance from the teacher
- Made me nervous because professor said exam was going to be harder
- Personal interaction affected attitude

**Outside Circumstances**
- Did better than I thought on exam so a bit more positive
- Didn't do well as I thought on exam
- Happy with how I did on exam
- Got a pretty bad grade
- Different view if didn't have exam
- Came into class not wanting to take it a third time
- Previous time taking class I had trouble
- Never liked math so was a bit nervous
- Apprehensive going in once I saw the syllabus and saw tests would be online
- A lot of people tell me statistics is hard
Theme 1: Classroom Structure. As with comprehension, classroom structure was a top emerging theme. For attitude, classroom structure included the codes of a flipped classroom, content, in person classes, and video lectures. Additional codes included “the material getting hard” and “group work.” Students in the treatment group shared how various aspects of the classroom structure were “frustrating”. “What affected my attitude more was that most of those in class activities were as a group, instead of individually. I tend to lose focus [with groups] so it's kind of frustrating,” said one student. Another student in the treatment group said, “There is too much to try to learn. It is too hard. It is super frustrating. I do not want to be in school more than 16 weeks but maybe make statistics a 20-24 week [course]. There would be more time per concept.”

The classroom structure also affected the attitude for students in the control group. “Sometimes I would walk out of class feeling like it was kind of wasting my time. It would be a little bit frustrating” said one student. Another student shared, “I found that me and my group thought activities were actually rather discouraging. The difficulty level was increased. Everybody felt like not going to class because the activities made us feel bad because nobody knew what they were really doing.”

There were some aspects of the classroom structure that were beneficial according to the students. “I am definitely an in-person learner. Being in the classroom helps me feel like I know what I am doing and getting immediate feedback helps,” said a student from the treatment
group. “Good attitude would be the video lectures because after viewing them, I would feel better about the concept,” said one student from the control group. This student also shared that “the formula sheet helped take a little bit of the stress off.”

Theme 2: Instructor. As previously mentioned, the instructor is the person providing the in-class instruction. Codes that composed this theme include the attitude of the instructor, knowledge of the material, willingness to answer questions, instructor feedback, and ability to personally connect with the students. These codes emerged in interviews from students in the control group and treatment group. “After the test, [my attitude] went really down. After talking with [the instructor], I understood where I messed up and how I could get my grade backup. As a result, [my attitude] went back up,” said a student in the control group. A student from the treatment group talked about the instructor being “willing to stay after class a bit for homework questions” helping with their attitude. Another student in the treatment group shared that the instructor “didn’t make [them] feel stupid” and that comfort helped maintain their positive attitude.

Category: Outside Circumstances. This category examined the elements outside of the instructional process which affected student attitude. Exam performance was mentioned by approximately 70% of those interviewed and received more than 25 mentions. “Personal life”, “peers”, and “group work” were the most referenced codes used for this category. Additional codes included
“preconceived notions” in addition to “previous experiences and perceptions” with statistics and math.

Theme 1: Exam Performance. Interview participants related their exam performance to their current attitude. While the quantitative SATS-36 was administered prior to the exam, interviews were not able to be conducted until afterwards. Performance on the exam, either doing well or not doing so well, was fresh on the minds of the participants being interviewed and was reflected in their responses. The strong response linking attitude to exam performance was unexpected. There was some attempt to determine if attitude responses would have been different if interviews were conducted prior to the exam. Students in the treatment group and control group highlight how the exam performance was a factor in their attitude.

One student said, “I honestly thought I was going to do worse on my exam. I did better than what I had hoped for. I guess [my attitude is] a little bit more positive right now. I was already very frustrated with the course. Taking the exam kind of relieved some of that. So yes, if you had interviewed me before [the exam], I probably would have had a more negative response.” Another student shared, “Before the exam I would have said moderate [attitude] but afterwards, I would say lower because I got a pretty bad grade.” The correlation between attitude and exam performance can further be summed up by this quote from a student, “I did
not do as well as I thought I would have done. That might be the only thing about the unit that has negatively impacted me or my mindset.”

Theme 2: Previous Experiences. The other emerging theme for outside circumstances was previous experiences. This theme is defined as having preconceived attitudes toward statistics based on previous experiences prior to this course. Codes included participants referring to not being a math person and having the want to succeed. “I am probably still a more negative attitude toward statistics. [Math is] not really my thing and I’ve always been not great at it. It's always been hard for me to follow along,” said a student in the control group.

Other students talked about how their experiences taking the class previously affected their initial attitude. “I think that I definitely came into this class like I was not taking this a third time. I am going to work my butt off. I had a goal to make sure I did it right, this time,” said a student in the treatment group. Another student in the treatment group mentioned, “I honestly think that [having taken the class before] did help ... I knew what I struggled with before.”

Students also shared how their overall attitude affected their attitude in the unit. “I am trying to stay positive because it is easy to just give up or not give everything your best effort. But I'm not really like that at all, actually. If I end up with a B, I just want to end up knowing I tried my best,” said one student in the treatment group. Another student
commented, “If I go into it, with at least a neutral attitude I feel like I might do a little bit better.”

The Effect of Context on Attitude. Similar to comprehension, the interviews included a direct question on how contexts used during the in-class activities affected attitude. One student in the control group provided an overall tone about the difficulty in adjusting attitude by stating, “Nothing too much [could have been done to sway my attitude]. That is on me. I think it is because I struggled for so many years.” This section links to research question two. Themes emerging under this category included the relevant context affecting attitude by being interesting, relatable to a career or to the real world, and being interpreted as useful.

Theme 1: Interesting. Interesting is defined as sparking engagement beyond what is encountered in a typical lesson. Codes for this theme include engaging, exciting, fun, and keeps attention. Even with the context not being tailored to the students in the control group, students found certain contexts used during the in-class activities as interesting.

One student in the treatment group commented how the relevant context “definitely made me pay more attention.” Even with the context helping with attention, it did not seem to be enough to fully affect learning. “Now I am not saying it certainly helps with understanding, but I can see the relation to what we’ve been learning,” said another student in the treatment group.

Students in the control group surprisingly also found the context in their unit interesting. “I have not had a math class that has these kinds of younger topics. It is always like the weird grocery store examples such as how many
watermelons in a grocery corner. These are actually like really nice college kid examples,” said a student in the control group. Another student said, “There are some things that we did in class that were more interesting, such as the babies one even though I do not work with babies. It is just interesting to think about. The runway on the sea with the paper airplanes was pretty interesting as well.”

Another student provided additional context by stating, “It is interesting because now, I know how to read different kinds of data, and I find that interesting to me, because that helps me better comprehend it and real-world situations when I fight with people about statistics, so I guess I kept a regular positive attitude about it.”

Another student from the control group summarized the connection between interest and attitude best by stating, “Assignments that spark happiness, or spark attention, seem to leave me with a better attitude.”

Theme 2: Relatable (Real-World). Relatable for attitude includes real-world applications, relating to careers, and the context seen as being practical. There was a distinction made between the control and treatment groups on relatability of the context used. There was no surprise that students in the treatment group commented on how the relevant context provided a better understanding of how probability is applicable in the future. “I had no idea where I would apply [the probability material] and the in class activities kind of helped with that.” said one student. Another student said, “I appreciate that the context puts the information into a perspective that the student can understand.”

Students in the control group shared how they were trying to connect the content to relevant experiences. One student said, “Whenever we were going
through these contexts, I tried to see how this is useful in real life.” Another
shared, “I feel like when I go back to work and I see people pick up candy, I will
wonder what the probability is that they would pick Skittles over M&Ms. I feel
like that made me interested.”

Students in the control group also provided insight as to what would make
the context relatable to them. “Even though I know the statistics itself is
important, it doesn't seem as relevant to me because I know I am never actually
going to do anything like that with earthworms. But if it is in the context of
something that I might actually do in business, then it is more important because I
could see how it could be implemented in my actual job.”

Theme 3: Useful. Useful is defined as noting the use of statistics material
in everyday occurrences and instances. Codes under this theme included helpful,
useful, and positive. This theme emerged mostly from students in the treatment
group. One student explained that “useful is something that I am going to be
using on the very basic levels.” One student provided more information about the
impact of the context being useful. “I recognized the business field [statistics] as
being highly useful. I would like to be able to know what people tell me when they
use statistics. However, it is not useful enough to weigh out the opportunity cost of
the rest of the things that I do in my week. I currently pursue money over the
slight amount of change I would make in learning statistics,” said a student in the
treatment group. Another treatment group student provided gratitude for useful
context in saying that “it is helpful to see something relevant because that's kind
of what I had been looking for, to be able to come in and compare what I am learning in class, to see if I can gain a better understanding through it.”

**Connecting Attitude and Comprehension.** Research suggests and supports a connection between comprehension and attitude (Hay et al., 2015; Ramirez et al., 2012; Tempelaar et al., 2007). As depicted above, attitude shifted based on the perceived comprehension of the student based on exam performance. One student in the treatment group defines how a class being fun can increase comprehension. “It was fun. The better time I have in class, the more excited I am to go back even if it is statistics. If I am in a better mood when I enter the door, I usually learn better.” This was reiterated by a student in the control group. “The more interested you are in learning something, then the more you're going to be highly focused and want to learn more about the topic. If you're not interested, then you probably just want to know the bare basics and don't want to go further.”

A student in the control group provided their explanation on the link between comprehension and attitude in this quote: “I would say that my comprehension relates more to my interest level and the subject, but I would not say that the relatability of something does not have a role in my interest. My comprehension plays a bigger part in how interested I am in the subject than the relatability of something plays into my comprehension of the subject and my interest in it. I would say that whenever you have an attitude in your head that you're going to be good about something, like a positive attitude about it, you are more likely to believe that you can do it. If you have a negative attitude about something you are more likely to believe that you can't do something.
Either one of those affects your comprehension. If you have a bad comprehension and if you have a negative attitude about it, you're not as interested in it.”

Summary

Using a mixed methods approach, an attempt was made to understand the impact of relevant context on student comprehension and attitude. With regards to comprehension, the treatment group did slightly outperform the control group both in change in pretest score and in overall posttest score. However, neither of these respective differences was large enough to suggest statistical significance ($p=0.94$ and $p=0.41$). The results did not change meaningfully when only analyzing students who attended a certain number of classes. Thus, the use of relevant context alone was not enough to create a significant difference in student comprehension of probability concepts.

In the quantitative analysis of attitude, it was found that student attitude actually became less positive over the course of the intervention in both the treatment and control groups. This is likely attributed to the overall difficulty of the unit. However, the differences between treatment and control groups yielded $p$-values well above the threshold for significance across all six of the attitude components that were studied. Thus, the use of relevant context alone did not create a significant difference in student attitude toward statistics. Follow-up analysis separating the student data into groups by instructor pointed to some potentially significant differences in student Affect and Interest toward statistics that may be attributed to the individual instructor characteristics.

Finally, a qualitative analysis using student interviews and focus groups sought to understand the role that relevant context played in student comprehension and attitude toward statistics. Additionally, these findings helped to characterize driving influences in
student comprehension and attitude. While findings on the impact of relevant context were somewhat mixed for both comprehension and attitude, other factors emerged that played a larger role in shaping student learning outcomes. Class structure, the amount of material, and the instructor were the driving factors that affected student comprehension. Regarding attitude, class structure and the instructor were the two primary factors within the classroom that impacted students. Finally, the outside circumstances of exam performance and previous math experiences also played a role in determining student attitude toward statistics.
Chapter 5

Discussion

Introduction

Under the existing education system, many students are exposed to only a single course in statistics. A significant percentage of students emerge from this course without a strong understanding of key concepts (DelMas et al., 2007). This lack of comprehension, combined with a negative attitude toward statistics, make it unlikely that these students will choose to use statistical reasoning in their futures (Schau & Emmioglu, 2012). Prior research shows that a positive shift in student attitude toward statistics frequently coincides with deliberate changes in classroom structure and instruction (Schau & Emmioglu, 2012). Through the research of different instructional strategies, educators can be equipped with the tools necessary to improve the comprehension and attitude of their students.

Within a statistics course, the concept of probability represents a significant hurdle to student success (Cobb, 2015; Garfield & Ben-Zvi, 2007). A 10-lesson intervention at an urban midwestern university during the probability unit of an introductory statistics course was implemented in an attempt to overcome this challenge. Two hypotheses were tested regarding the effectiveness of using relevant context as a strategy to improve student comprehension and attitude toward statistics. Additionally, follow up interviews and focus groups sought to understand the factors that influenced student comprehension and attitude. These findings may provide insight into future research surrounding the use of relevant context as an instructional strategy.
Summary of Findings

Comprehension

Two separate measures were used to determine student change in comprehension: the change between pretest and posttest scores and the overall posttest score. These two measures were first analyzed separately using two-sample \( t \)-tests, and then the results from those tests were considered together when attempting to answer the first research question: To what extent does using relevant context within situated experiences affect student comprehension of probability concepts in a collegiate introductory statistics course?

The same 10 questions from the pretest were embedded in the posttest with the gain between pretest and posttest scores calculated for each student. The average change in these pretest to posttest scores for treatment and control groups was compared using a two-sample \( t \)-test. Although the treatment group had a higher average gain than the control group, no significant difference was found in the change from pretest to posttest scores between the two groups.

The second measure that was used to determine student change in comprehension was the overall posttest scores. The average posttest scores of the treatment and control groups were also compared using a two-sample \( t \)-test. Again, the treatment group did have a higher average than the control group, but no significant difference was found between the posttest scores of the two groups.

Because some students did not regularly attend class and therefore could not benefit from the intervention activities, \( t \)-tests were also conducted for participants in the treatment and control groups who were present for a majority of the intervention
activities. For students with higher than 50% attendance, there were still no significant differences found for either measure of comprehension. Based on the cumulative findings of these results, using relevant context within situated experiences did not appear to significantly affect student comprehension of probability concepts in a collegiate introductory statistics course.

**Attitude**

As specified by the *Survey of Attitudes Toward Statistics (SATS-36)*, the six components under study were Affect, Cognitive Competence, Value, Difficulty, Interest, and Effort. The change in each of the six attitude components was calculated for every student, based on *SATS-36* scores before and after the probability unit. These changes were analyzed to answer the second research question: To what extent does using relevant context within situated experiences affect student attitude toward statistics?

Before addressing the research question, it is worth noting that the average change in attitude was negative for all six components of students in the treatment group and negative for five of the six components for students in the control group, with the component of Difficulty remaining unchanged. While these attitude components did decrease slightly, all attitude component averages after the intervention were still either positive or neutral. In other words, attitude averages were less positive, rather than negative, upon the conclusion of the probability unit. Because this decrease occurred in both the treatment and control group, it is likely that this was due to factors other than the use of relevant context, such as the challenging nature of the probability content.

When comparing the attitude scores of students in the treatment group versus the control group, the difference in attitude change was not found to be significant for any of
the six components. Even though attitude decreased on average for both the treatment group and control group, the rate of decrease was similar across both groups. In response to the second research question, the use of relevant context during this intervention was not enough to create a significant difference in student attitude toward statistics.

After comparing the treatment group to the control group to answer the second research question, a more granular look was taken at the attitude data, separating it into three groups based on the three different instructors involved. While changes in group attitudes remained consistent across many of the six components, two key differences emerged when separating the data by instructor. Students in the treatment group section taught by the instructor with 17 years of experience emerged with substantially greater positive outcomes in Affect and Interest toward statistics than those in either the second treatment group section or the control group, taught by instructors with three years of experience. This finding is consistent with prior research such as Di Martino & Zan (2010) and Hidi & Renninger (2006), and provides further evidence for the role of the instructor in shaping student attitude toward statistics. Further support for the importance of instructor was uncovered in the qualitative interviews conducted after the intervention.

**Qualitative Follow-Up**

Sixteen students participated in 10 interviews and two focus groups to better understand the factors affecting student comprehension and attitude. The one-on-one interviews and focus groups were transcribed and coded using a deductive coding approach. Using this approach, themes emerged to answer the third research question: In what ways do students describe the factors that affected their comprehension of probability and their attitude toward statistics based on the probability unit?
Transcripts were uploaded into NVivo, a qualitative data analysis software, to assist with the coding process. The initial codes were determined based on the in vivo coding method. An iterative process was used to finalize the list of codes before determining the emerging themes.

*Class Structure, Amount of Material, and Instructor* were the themes that emerged as elements of the instructional process that affected comprehension. *Schedule* was the singular theme from outside circumstances that factored into comprehension. Schedule affected the amount of time students were able to devote to learning course content. Three of the most frequent codes listed under the theme of schedule were work schedules, family, and home responsibilities. When asked about how the use of relevant context affected comprehension, those interviewed iterated that some aspects of the context were “interesting”. Therefore, interest had a perceived effect on student comprehension, but the relevant context did not emerge as being an effective tool in improving student comprehension as highlighted in the qualitative portion of the study.

*Class Structure and Instructor* were the themes that emerged as elements of the instructional process that affected attitude. *Exam Performance* and *Previous Experiences* were the themes for outside circumstances, factors beyond the elements of the instructional process that affected attitude. As previously mentioned, there was no significant effect from the use of relevant context on attitude based on the quantitative portion. Despite these findings, some of the students still mentioned that the relevant context was “interesting”, “relatable”, and “useful” during the qualitative portion.

Although no statistically significant link was found between the use of relevant context on either comprehension or attitude, the qualitative results indicated that there
were other factors that students believed may have affected these learning outcomes. The classroom structure, the amount of material covered, the instructor, student schedules, exam performance, and previous experiences were the main influencing factors shaping student views on their learning outcomes in the probability unit.

**Limitations**

Over the course of the intervention, several unavoidable factors may have limited the applicability of the findings to other settings. Two such limitations concern timing with respect to the COVID-19 pandemic. First, this study was implemented in the first semester of a return to in-person learning following a year and a half of virtual instruction. This prior experience with virtual learning could have influenced student comprehension and attitude toward statistics in ways that do not generalize to other groups of students. Additionally, enrollment in the in-person sections of the course was substantially less than in a typical year due to the continued offering of virtual sections under the pandemic. Rather than the usual four sections of the course, only three sections were offered during the semester under study, with a lower number of students in each section.

In addition to these practical issues driven by the pandemic, two other realities of university scheduling created limitations with regards to the desired similarity between treatment and control groups. Due to constraints in staffing, all three sections were taught by different instructors with different levels of experience, which made it impossible to balance the influence of the instructor on student comprehension and attitude. Finally, the design did not incorporate random assignment to sections, but rather used a quasi-experimental design. Although the quantitative measures did assess for initial differences
in comprehension and attitude between treatment and control groups, it is possible that the sections may not be considered equivalent due to a variety of other factors. One course section was offered in the evening, which traditionally attracts students with additional commitments such as full-time jobs and families. Additionally, students could have placed themselves into any section of the course based on familiarity with the instructor. These factors, and others not anticipated, could have created class sections that may not be perfectly equivalent to one another.

**Delimitations**

Due to the scope of work required in designing relevant contextual activities from scratch, along with the parallel generic lessons, the decision was made to implement this intervention in a single six-week unit. Research suggests that it may take time to change student attitude toward a subject (Bateiha et al., 2020; Schau & Emmioglu, 2012), so the choice to limit the study to a single unit may not have revealed the true impact created by the relevant context. Furthermore, the response of students to the intervention may also be influenced by the difficulty of the probability concepts covered within the unit, rather than the relevant context itself.

All quantitative measures for comprehension and attitude were administered remotely, as in earlier semesters during the pandemic. While this decision improved the convenience of managing the schedule for both students and instructor, it does limit the ability to determine whether the results of each student are authentic. An online proctoring system was used during the comprehension unit assessment, but it cannot be certain that none of the students under study engaged in academic dishonesty while testing. Also, because researchers did not have control over the testing environment for
students during the exams and attitude surveys, it is possible that other factors such as external distractions may have influenced student responses.

Lastly, the choice of contexts used were not relevant to each individual student in the treatment groups. Contexts were selected based on overall interest level for all students, and individual students in the treatment groups may not have found certain topics relevant as intended. Conversely, no attempt was made to ensure that the topics chosen for the generic context were minimally relevant to students in the control group. Rather than deliberately choosing topics rated the lowest on the interest inventory, parallel lessons were created to those given to the treatment groups, selecting topics that seemed less relevant than those in the treatment group. Because of this method of context selection, it was not guaranteed that there would be a difference in relevance for each intervention lesson. The difference in relevance between treatment and control groups was only ascertained after analyzing exit slip data.

Implications for Practice

Comprehension

As depicted in the Chapter 4 quantitative results, relevant context did not have a significant impact on comprehension for the students in the treatment group. Based on these findings, choosing a context for an activity to highlight a particular concept does not necessarily need to be tied to the personal interests of the students. It could be better to select a context based on the ease of the instructor to connect that context to the content of the lesson. When asked whether the relevant context activities were helpful in learning the content, one of the treatment group students stated “Not so much. I feel like some of them were, but some of them just seemed silly...The real world examples are
really helpful when you know what all the different pieces mean and you just put them together and see if you can apply it. But if you don’t have the [mathematical] foundation, then the real world is just confusing and silly sometimes.” Establishing a strong connection between context and content would potentially lead to fewer students feeling like their classwork or activities were not useful.

However, before an instructor defaults to traditional contexts of dice, decks of cards, and balls in urns, the qualitative follow-up provided some additional consideration for instructors as they add context to content. First, despite the study results, using bland contexts may not be as engaging to students as those that do contain more personal relevance. Students mentioned that relevant contexts were “interesting”, “useful”, and “helpful” as part of the instructional process. One student in the control group stated, “I think it is easier if you are learning about something that you already know about or if it is geared towards something that will be important in my job.” This is also supported by the findings of Neuman et al. (2013) and Garfield & Ben-Zvi (2009) which found that real-world datasets were valuable for the comprehension of some students. While an instructor may not need to tailor the context of the lesson to the personal interests of each individual student, finding examples that are engaging or relatable to life experiences or future careers could help students in their comprehension of topics.

Some students who were interviewed identified a disconnect between the relevant context found in the in-class activities and the homework assignments and tests, which had not been modified to enhance relevance. According to a student in the treatment group, “It was helpful to see relevant context in the class activities because that was what I had been looking to see. But when I saw the homework, it was a little bit different. The
context was not relevant. It was back to Johnny rolling dice.” As a result of this discrepancy, some students did not see the value of the in-class activities, instead viewing them as childish or a distraction. If consistent, contextual connections were established between lectures, homework, and assessments, students may have found more meaning and value their overall experience, aiding in their comprehension of topics.

As one student in the treatment group stated, “[using relevant context] is definitely moving in the right direction. It has improved my comprehension over how I would have done in a normal 100% old school math style textbook.” This same student went on to say that “it was more helpful to have a real-world example with something that is abstract and a hard concept to normally comprehend.” While the results indicated that relevant context may not be a cure-all to help students comprehend probability topics, finding an interesting connection to the real-world is still a valuable tool that teachers can employ.

**Attitude**

According to the analysis of the change in attitude data, the use of relevant context did not create a significant difference in any of the six attitude components between the treatment and control groups. Similarly to the implications regarding comprehension, this suggests that a teacher does not need to go out of their way to seek topics that are personally relevant to student interests. Again, it may be more advantageous to select contexts that naturally showcase a given topic, rather than attempting to force a particular context into a lesson. This may capture student interest, even if the topic is not something they are already familiar. As one student from the treatment group observed, “If the context is something that I might actually be doing in
the future, then it is more important because I recognize at least that I could use this and I could see how it could be implemented in my actual job.” If the instructor can draw student attention to relevant applications of the content, there is opportunity for the students to find increased value in the lesson.

Although no significant differences in comprehension or attitude were found between the treatment and control groups, subsequent analysis detected a difference in attitude change when separating all three sections by instructor. In the course taught by the most experienced instructor, a substantially higher percentage of students emerged with a positive shift in their Affect and Interest toward Statistics. In other words, students in the section with the most experienced instructor tended to like statistics even more and find it more interesting in comparison to those in the other sections of the course. This quantitative finding is consistent with prior research, which suggests that the instructor can have a major influence on student learning outcomes (Di Martino & Zan, 2010; Hattie, 2015; Hay et al., 2015).

Finally, findings in the qualitative portion reaffirm the known link between comprehension and attitude. Higher performance in statistics is associated with a more positive attitude, while struggles in comprehension can lead to a more negative attitude and potential avoidance of statistics (Di Martino & Zan, 2010; Ramirez et al., 2012). Due to logistical constraints, the qualitative interviews were not conducted until after students completed the unit posttest. A number of interviewees described their performance on the unit exam as a driving force on whether they held a positive or negative attitude toward statistics. As one student stated, “In my mind what tells me if I am doing good or feeling good is my exam score and keeping my grade up.” As instructors work to inspire a more
positive attitude toward statistics for their students, it is essential to consider the role that comprehension plays.

**Recommendations for Future Research**

**Pedagogy**

While the results did not reveal a significant link between relevant context and improvements in comprehension or attitude, further research is warranted regarding the scope and the length of a relevant context-based intervention. When an instructor is successful in inspiring major changes in student attitude, it is frequently accompanied by substantial changes to the instructional process (Schau & Emmioglu, 2012). Attitude tends to be somewhat stable (Gal et al., 1997), and is more likely to improve when students encounter a pivotal change in their learning experience (Di Martino & Zan, 2010). The study took place over a six-week unit with 10 activities. Future studies should consider implementing an intervention that sustains over an entire course, as this could lead to stronger differences in student learning outcomes.

Further research may also be warranted in providing some sort of “interesting” context as a part of the normal practice of teaching, intertwined into every aspect of the instructional process including lectures, in-class activities, homework, and assessments. As noted in the implications for practice, students who were interviewed described the disconnect between the types of contextual examples done in class and those used in assignments and assessments. The discrepancy between bland and relevant contexts in various parts of the course may have caused some students difficulty in their comprehension of course content. Consistency in the implementation of context across
different aspects of the course may improve clarity and support student comprehension and is a recommendation for future study.

Finally, due to logistical constraints, this intervention utilized three separate instructors. Given the strong influence of a teacher on learning outcomes (Hay et al., 2015), this difference in instructor may have made it more difficult to ascertain the exact impact of the relevant context. If a similar study were to be conducted using the same instructor teaching multiple sections of the course, it would become even clearer how much of an impact the relevance had on student comprehension and attitude.

**Relevance**

Although no significant difference was found in comprehension or attitude based on the use of lesson-specific, relevant contexts, an area of future research could be to see if repeated use of one or a few select contexts for multiple lessons would make an impact on the learning outcomes. Naidu & Sanford (2017) described their use of Russian Roulette over the course of many weeks to illustrate various probability topics. While no rigorous data was collected, the researchers believed that the single context aided in comprehension and was an example of interest to many of their students. Wathen & Rhew (2019) also did not formally conduct a research study, but described their use of baseball statistics as motivating student interest in statistics over a number of lessons. Similarly, Lyford et al. (2018) also did not collect quantitative data, but used a game called Camel Up to teach probability concepts and expected value. Their end of semester evaluations had positive student comments about the content and interest level for the class. Conducting a formal research study to determine the impact of the repeated use of a
small number of contexts on student comprehension and attitude toward statistics could be an area of future research.

Despite that the quantitative results showed no statistical significance, the qualitative results did suggest that some students benefited from the use of relevant contexts. A future study could use a similar methodology but increase the personal relevance for students in the treatment group by creating multiple activities for each lesson, each with a different context but covering the same content. Then students could self-select the context that would be most relevant to them for each lesson. Sports was a topic of interest identified by a majority of students and that context was deemed to be sufficiently relevant when creating activities. However, a student from the treatment group mentioned, “I'm not a big sports person. I understand very little about sports, so for some of the [activities] that were related to that, I was a bit confused...[and it] was a little harder for me to grasp.” For a student like this, having an option to select a more personally relevant context could have made for a more meaningful in-class experience. Further support for this recommendation came during an interview, when a different treatment group student explained, “I like the concept of picking things people are interested in, because I have been drawn in by some of [the contexts] and been like ‘oh that's an interesting topic.’” Prior research also suggests that student choice in contexts could improve comprehension. Walkington & Bernacki (2018) postulated, “allowing students ownership and choice [in the contexts of their lessons] may increase cognitive engagement.” Therefore it is possible that adapting the in-class activities to create context options for students may have an impact on student learning outcomes.


**School Setting**

Qualitative findings indicated that previous experiences factored into student attitude toward statistics. Participating students in this introductory statistics course ranged in age, gender, education level, citizenship, and choice of major. Research shows that diverse identities and differing cultures are factors that affect attitudes and “beliefs toward mathematics” (Gal et al., 1997, p. 40). One factor that emerged during the interviews was the impact of age. During an interview, one of the treatment group students stated, “I’m older than most of those students, so some of the things I understood when they didn’t, and then some of the things they understood, and I didn’t.” A recommendation would be to conduct this intervention in a school setting with younger students, who fall into a narrower age range and thus, background of experiences. As students develop preconceived attitudes from their experiences in and out of the class environment (Gal et al., 1997), this school setting would remove a possible external factor that affects attitude and comprehension as the students would be younger. This school setting would provide a study population with students at an earlier age, making it less likely that they would have a strong preconceived attitude toward statistics or mathematics in general.

Students during the qualitative portion also mentioned that the amount of information during the unit was an element of the instructional process which affected their comprehension. The participants spoke about how the content builds upon itself, and that there was not sufficient time to digest and comprehend one topic before moving on to the next. In order to remedy this factor, another recommendation is to offer the relevant context intervention over a longer period to ascertain if it would then have a
significant effect on comprehension or attitude. This would also allow for more consistent contextual connections, as discussed previously, and it could allow the relevant context to be integrated into multiple aspects of the instructional process. Research shows that experiencing data in different scenarios, situations, and forms increases the ability for the students to comprehend (Hay et al., 2015). A future study could determine if consistency and reinforcement of relevant content over a longer period would factor into attitude improvement and increase in comprehension.

Conclusion

In examining the use of relevant context as an instructional strategy, this study sought to add to the existing knowledge base regarding student comprehension and attitude toward statistics. Based on the quantitative findings, the use of relevant context failed to create a statistically significant difference in either comprehension or attitude toward statistics for the students under study. Subsequent qualitative findings point toward factors such as instructor characteristics, classroom structure, and previous math experiences that may have played a larger role in shaping student comprehension and attitude.

While the use of relevant context did not create a significant difference in student learning outcomes, this is not to say that the pursuit of relevance for students is fruitless. Qualitative findings did uncover enhanced interest in statistics from a portion of those interviewed. However, if the primary goal of the instructor is to improve student comprehension and attitude toward statistics, both quantitative and qualitative findings suggest that the answer to this task may be elsewhere. The field of statistics poses
significant challenges for many students, and future research will be needed to uncover best practices for success in this discipline.
References


https://doi.org/https://doi.org/10.1080/10691898.2020.1765710


https://doi.org/10.1080/10691898.2019.1669508


https://doi.org/https://doi.org/10.7275/bj1p-ts64


https://doi.org/10.1080/10691898.2002.11910679


https://doi.org/10.1080/10691898.2009.11889515


https://doi.org/10.2307/1403775


Schau, C. (2021, May 22). *Register and Request SATS Form* [Personal communication].


https://higherlogicdownload.s3.amazonaws.com/AMSTAT/56c109df-f1a1-45e2-bb2b-a6880c16f76a/UploadedImages/v7n1.html#Perusing


https://doi.org/10.1007/s11858-018-0967-8
Informed Consent for Participation in Research Activities

THE IMPACT ON STUDENT COMPREHENSION AND ATTITUDE USING RELEVANT CONTEXT WITHIN SITUATED EXPERIENCES IN A COLLEGIATE INTRODUCTORY STATISTICS UNIT ON PROBABILITY

Participant ______________________________   HSC Approval Number ___________________

Principal Investigator: ____________________________   PI’s Phone Number: _____________
Co-Investigators: _________________________________

Summary of the Study

This is a brief description of the project:

You are invited to participate in a research study about the impact of relevant contextual examples on student comprehension and attitude. The study will be conducted on the probability unit (Chapter 5) in this course. This unit lasts approximately six weeks. Your participation in this study is voluntary. If you agree to be part of the research study, you will be asked to complete a pre- and post- survey centered on your attitude towards statistics. You also can also participate in either a focus group or an individual interview once the probability unit has ended. All students, whether participating in the study or not, will complete a pre- and post- assessment and participate in group in-class activities as a normal part of the coursework. The foreseeable risks of participation include a loss of confidentiality as the researchers will match your pre- and post- assessment scores as part of the study. While there are no expected direct benefits to you, benefits of the research include adding to the collective knowledge about the use of relevant contextual examples in collegiate-level courses. Because the research takes place within the normal scope of your coursework, there are no alternative procedures should you choose to not participate.

1. You are invited to participate in a research study conducted by ____________________________ The purpose of this research is to determine the impact that
using relevant contextual examples have on student comprehension and student attitude in an introductory statistics course.

2. a) Your participation will involve the following (in chronological order)
   Completing an interest survey at the start of the course (as a normal part of introductory coursework)
   Taking the SATS-36 pre-survey before the probability unit begins
   Participating in the normal classwork for this course, including but not limited to:
   o taking a Chapter 5 pretest
   o participating in in-class activities
   o providing feedback on exit cards at the end of each activity
   Taking the SATS-36 post-survey after the last in-class activity but before the posttest
   Taking both portions of the Chapter 5 posttest
   Optionally, participating in a recorded Zoom interview or focus group

You will be assigned to small groups of 3-4 students at your instructor’s discretion for in-class activities. These activities will take place twice a week for the duration of the probability unit (Chapter 5).

Students will self-select to participate in the interviews/focus groups. If more students volunteer than is needed, the researchers will purposefully select students with varying scores on the attitude surveys and/or comprehension assessments to ensure a variety of perspectives were captured.

Approximately 135 students may be involved in this research at the University of Missouri-St. Louis.

b) The amount of time involved in your participation will be approximately 45 minutes above the normal time requirements for your course to take the pre- and post- attitudinal survey unless you also volunteer to participate in a recorded zoom focus group or interview. In this case, the additional time involved will be somewhere between 30-90 minutes and you will receive extra credit in the course for this additional time. Should you wish to participate in the interview and/or focus group, but were unable to do so, you will have the opportunity to earn those same extra credit points by writing an essay about your experiences.

3. There are no known risks associated with this research other than the loss of confidentiality. The researchers will minimize the risk associated with the loss of confidentiality by replacing your names with unique identifiers in our databases, by giving pseudonyms to interview or focus group participants, by destroying all recordings after the conclusion of the study, and by keeping all documentation on password-protected devices during the study and destroying them at the conclusion of the study. Information collected in this project may be shared with other researchers, but the researchers will not share any information that could identify you.
4. There are no direct benefits to you for your participation in this study. Your participation will not lead to any additional credit in the course, aside from the extra credit opportunity described in 2b.

5. Your participation is voluntary and you may choose not to participate in this research study or withdraw your consent at any time. You will NOT be penalized in any way should you choose not to participate or withdraw. Non-participation will not affect your standing in the course. While the pre/post assessments are part of the class curriculum and everyone participating in the research or not will be taking them, consenting to participate in the study allows the use of your data for research purposes.

6. We will do everything we can to protect your privacy. As part of this effort, your identity will not be revealed in any publication that may result from this study. In rare instances, a researcher's study must undergo an audit or program evaluation by an oversight agency (such as the Office for Human Research Protection) that would lead to disclosure of your data as well as any other information collected by the researcher.

7. If you have any questions or concerns regarding this study, or if any problems arise, you may call the Investigator. [REDACTED]. You may also ask questions or state concerns regarding your rights as a research participant to the Office of Research, at 516-5897.

I have read this consent form and have been given the opportunity to ask questions. I will also be given a copy of this consent form for my records. I hereby consent to my participation in the research described above.

Participant's Signature _______________________________ Date _______________________________

Signature of Investigator or Designee _______________________________ Date _______________________________
Volunteer / Extra Credit Opportunity

As part of the educational research being conducted this semester, one of the researchers would like to conduct zoom interviews and/or hold a zoom focus group with students to discuss their experiences during the Exam 2 material within a few days after Exam 2 has been completed. If you are selected, your focus group or interview will take place in one session and last between 30-45 minutes for the one-on-one interview and 75-90 minutes for a focus group with three (3) other students.

Students who are selected for either an interview or a focus group will earn 10 points of extra credit in the course after meeting with . Students who are not selected or do not feel comfortable participating but still wish to earn extra credit may do so by writing an essay instead and emailing it to no later than March 30th. Directions for this essay will be posted on Canvas after Exam 2 has been completed.

The focus group and interviews will be conducted via Zoom and will be the only person who will be able to know who made certain statements. will transcribe the conversation and before any of the focus group or interview data is released, he will give each participant a pseudonym. will provide your instructor the list of students who earned extra credit, not noting who participated in the interview/focus group and who wrote an essay.

I would like to participate in (check the appropriate choice)

- A one-on-one Interview via Zoom
  (Interview slots available at various times between March 16-28.)
- A focus group with three (3) other students via Zoom
  (Focus Group conducted 6-7:30pm on March 21, March 23, or March 30)
- Either an interview or a focus group
- Neither the interview nor the focus group

For those students who are willing to participate, please provide an email address that can be passed along to so that he can reach out to you if you happen to be selected.

Email address where I can be most easily reached
# Student Interest Inventory

Broad Topics - Rate your interest in experiencing in-class examples during the probability unit of your statistics class centered on the seven broad categories: Sports & Activities, Health, School & Career, Politics & Social Justice, Leisure Activities & Hobbies, Technology, and Finance & Business

* Required

1. How interested would you be in using examples from the following topics to help understand the concepts in this course? *

   *Mark only one oval per row.*

<table>
<thead>
<tr>
<th></th>
<th>Extremely Uninterested</th>
<th>Somewhat Uninterested</th>
<th>Ambivalent / Neutral</th>
<th>Somewhat Interested</th>
<th>Extremely Interested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sports &amp; Activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School &amp; Career</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Politics &amp; Social Justice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leisure Activities &amp; Hobbies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finance &amp; Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Specific Contexts*

Please rate their interest on a number of specific contexts related to each of the broad categories below.
2. Please rate your interest in these specific contexts related to Sports & Activities: *

*Mark only one oval per row.*

<table>
<thead>
<tr>
<th></th>
<th>Extremely Uninterested</th>
<th>Somewhat Uninterested</th>
<th>Ambivalent / Neutral</th>
<th>Somewhat Interested</th>
<th>Extremely Interested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectator sports</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(including Cardinals and Blues)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recreational sports</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaming / eSports</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exercise / Working Out</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gambling / Lotteries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Please rate your interest in these specific contexts related to Health: *

*Mark only one oval per row.*

<table>
<thead>
<tr>
<th></th>
<th>Extremely Uninterested</th>
<th>Somewhat Uninterested</th>
<th>Ambivalent / Neutral</th>
<th>Somewhat Interested</th>
<th>Extremely Interested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nutrition / Wellness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healthcare / Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental Health / Mental Illness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychology / Social Behaviors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contraception / STIs / Sex Issues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug / Alcohol Use or Abuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COVID-19 / Pandemics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4. Please rate your interest in these specific contexts related to School & Career: *

*Mark only one oval per row.*

<table>
<thead>
<tr>
<th>Issue / Context</th>
<th>Extremely Uninterested</th>
<th>Somewhat Uninterested</th>
<th>Ambivalent / Neutral</th>
<th>Somewhat Interested</th>
<th>Extremely Interested</th>
</tr>
</thead>
<tbody>
<tr>
<td>UMSL-Specific Issues / Student Affairs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Loans / College Affordability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online / Distance learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Service / Volunteering</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment Trends</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Please rate your interest in these specific contexts related to Politics & Social Justice: *

*Mark only one oval per row.*

<table>
<thead>
<tr>
<th>Context</th>
<th>Extremely Uninterested</th>
<th>Somewhat Uninterested</th>
<th>Ambivalent / Neutral</th>
<th>Somewhat Interested</th>
<th>Extremely Interested</th>
</tr>
</thead>
<tbody>
<tr>
<td>International Affairs / Current Events</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environmental Issues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Racism / Racial Inequities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Religious Discrimination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policing / Criminal Justice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immigration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender Discrimination / Sexism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGBTQIA+ Issues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Politics / Gerrymandering / Voter Issues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gun Violence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bullying</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food Insecurity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6. Please rate your interest in these specific contexts related to Leisure Activities & Hobbies: *

*Mark only one oval per row.

<table>
<thead>
<tr>
<th></th>
<th>Extremely Uninterested</th>
<th>Somewhat Uninterested</th>
<th>Ambivalent / Neutral</th>
<th>Somewhat Interested</th>
<th>Extremely Interested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traveling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Music</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shopping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ Carpentry / DIY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Art / Photography /</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theater</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Media</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conspiracy Theories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Dating</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooking / Baking / Food</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
7. Please rate your interest in these specific topics related to Technology: *

Mark only one oval per row.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Extremely Uninterested</th>
<th>Somewhat Uninterested</th>
<th>Ambivalent / Neutral</th>
<th>Somewhat Interested</th>
<th>Extremely Interested</th>
</tr>
</thead>
<tbody>
<tr>
<td>TV Binge Watching / Screen time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artificial Intelligence / Automation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internet Access</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space Exploration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. Please rate your interest in these specific topics related to Finance & Business: *

Mark only one oval per row.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Extremely Uninterested</th>
<th>Somewhat Uninterested</th>
<th>Ambivalent / Neutral</th>
<th>Somewhat Interested</th>
<th>Extremely Interested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economics / Stock Market</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing Prices</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insurance Rates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wealth Gap / Economic Inequality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum Wage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B.2

Interest Inventory Results and Narrative

*Interest Survey*

**Data Collection.** During the first week of the semester, prior to the probability unit under study, students were asked to fill out an interest survey so that the personal interests of treatment group students could be incorporated into the contexts used for the situated activities. Additionally, topics of high interest were purposefully avoided for control group activities, although many of the non-relevant contexts did not appear on the interest survey. Students were asked a series of questions on a Google Form, and they responded to each question by selecting one of 5 choices: “Extremely Uninterested,” “Somewhat Uninterested,” “Ambivalent / Neutral,” “Somewhat Interested,” and “Extremely Interested.” Questions were first asked about general topics of interest and then more detailed subtopics for each general topic were explored. Data collection was completely anonymous, but students were asked to identify which of the three course sections they were enrolled in.

**Intervention Fidelity.** Even though there were only 72 students enrolled in the three course sections, there were 83 responses to the interest survey. Therefore, some students filled out their responses more than once. However, because of the anonymous responses, it was impossible to eliminate duplicate results from individual students. Since this survey was only meant to guide the researchers in their planning of contexts and the exit slips would verify the amount of relevance for each particular context, the duplicate responses were not considered overly problematic.
Results. Students were first asked to rate their interest in the general categories of Sports & Activities, Health, School & Career, Politics & Social Justice, Leisure Activities & Hobbies, Tech, and Finance & Business. Results from these general categories can be seen in Figure B.2.1. All interest survey graphical results use the following legend:

![Legend](image)

**Figure B.2.1**

*Student Interest Survey Results of General Categories*

Each general category was then broken down into more specific subtopics, and students were asked to rate their interest in these subtopics using the same five Likert-response options. The subtopics for Sports & Activities were Spectator Sports (including local professional sports teams), Recreational Sports, Gaming / eSports, Exercise / Working Out, and Gambling / Lotteries (see Figure B.2.2).
Figure B.2.2

*Interest Survey Results for Subtopics under the General Category of Sports & Activities*

Please rate your interest in these specific contexts related to Sports & Activities:

![Bar chart showing interest levels for various subtopics under Sports & Activities.]

*Note.* To maintain anonymity of the research site, the specific names of nearby professional sports teams were removed from the graphic above. Participants did see the names as they were completing the survey.

The seven subtopics for Health included Nutrition / Wellness, Healthcare / Medicine, Mental Health / Mental Illness, Psychology / Social Behaviors, Contraception / STIs / Sex Issues, Drug / Alcohol Use or Abuse, and COVID-19 / Pandemics (see Figure B.2.3).

Figure B.2.3

*Interest Survey Results for Subtopics under the General Category of Health*

Please rate your interest in these specific contexts related to Health:

![Bar chart showing interest levels for various subtopics under Health.]

The five subtopics for the category of School & Career contained University-Specific Issues / Student Affairs, Student Loans / College Affordability, Online /
Distance Learning, Community Service / Volunteering, and Employment Trends (see Figure B.2.4).

**Figure B.2.4**

*Interest Survey Results for Subtopics under the General Category of School & Career*

Please rate your interest in these specific contexts related to School & Career:

![Bar chart showing interest in different subtopics](image)

*Note.* To preserve anonymity of the research site, the specific name of the university has been removed from the graphic.

There were 12 subtopics for students to rate under the main category of Politics & Social Justice. These included International Affairs / Current Events, Environmental Issues, Racism / Racial Inequalities, Religious Discrimination, Policing / Criminal Justice, Immigration, Gender Discrimination / Sexism, LGBTQIA+ Issues, Politics / Gerrymandering / Voter Issues, Gun Violence, Bullying, and Food Insecurity (see Figure B.2.5).
**Figure B.2.5**

*Interest Survey Results for Subtopics under the General Category of Politics & Social Justice*

The 11 subtopics for Leisure Activities & Hobbies were Reading, Traveling, Music, Pets, Shopping, Home Improvement / Carpentry / DIY, Art / Photography / Theater, Social Media, Conspiracy Theories, Online Dating, and Cooking / Baking / Food (see Figure B.2.6).
**Figure B.2.6**

*Interest Survey Results for Subtopics under the General Category of Leisure Activities & Hobbies*

Technology had only four subtopics: TV Binge Watching / Screen Time, Artificial Intelligence / Automation, Internet Access, and Space Exploration (see Figure B.2.7).

**Figure B.2.7**

*Interest Survey Results for Subtopics under the General Category of Technology*
The five subtopics for the final category of Finance & Business were Economics / Stock Market, Housing Prices, Insurance Rates, Wealth Gap / Economic Inequality, and Minimum Wage (see Figure B.2.8).

**Figure B.2.8**

*Interest Survey Results for Subtopics under the General Category of Finance & Business*

Based on these results and on the nature of the content to be taught, 10 contexts were selected for the treatment group which were considered personally relevant by a majority of students and 10 contexts for the control group which were not considered personally relevant.
APPENDIX B.3

Exit Slip Form and Exit Slip Data

Figure B.3.1

Exit Slip Form Administered After Each Activity

Name:______________________________________ Date:________________

Exit Slip

Context is what gives a mathematics or statistics problem meaning or describes the general setting in which the problem occurs.

Today’s activity centered on the context of ____________________________________.

A context is considered personally relevant if it meets any of these criteria:

- it relates to a prior experience
- it elicits positive emotional reactions
- it is associated with feelings of worth or value
- it is connected to important personal goals

Disregarding the mathematics and statistics from today, please rate how relevant today’s context was to you using this 5-point scale:

Not at all relevant 1 2 3 4 Highly relevant 5
### Exit Slip Data

**Table B.3.1**

*Exit Slip Data from Treatment Group 1*

<table>
<thead>
<tr>
<th>Student</th>
<th>Act 1</th>
<th>Act 2</th>
<th>Act 3</th>
<th>Act 4</th>
<th>Act 5</th>
<th>Act 6</th>
<th>Act 7</th>
<th>Act 8</th>
<th>Act 9</th>
<th>Act 10</th>
<th>Overall</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td>3.75</td>
<td>8</td>
</tr>
<tr>
<td>102</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>3.6</td>
<td>10</td>
</tr>
<tr>
<td>103</td>
<td>4.5</td>
<td>5</td>
<td></td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td></td>
<td></td>
<td>4</td>
<td>4.357</td>
<td>7</td>
</tr>
<tr>
<td>104</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td></td>
<td>3.667</td>
<td>9</td>
</tr>
<tr>
<td>105</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>3.25</td>
<td>4</td>
</tr>
<tr>
<td>106</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td></td>
<td>3.67</td>
<td>10</td>
</tr>
<tr>
<td>107</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>3.5</td>
<td>10</td>
</tr>
<tr>
<td>108</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3.625</td>
<td>8</td>
</tr>
<tr>
<td>109</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td>4.75</td>
<td>8</td>
</tr>
<tr>
<td>111</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>blank</td>
<td>4</td>
<td>3</td>
<td>3.778</td>
<td>10</td>
</tr>
<tr>
<td>112</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>3.7</td>
<td>10</td>
</tr>
<tr>
<td>113</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4.6</td>
<td>10</td>
</tr>
<tr>
<td>114</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4.5</td>
<td>8</td>
</tr>
<tr>
<td>115</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>116</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.333</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>117</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4.5</td>
<td>10</td>
</tr>
<tr>
<td>118</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.833</td>
<td>6</td>
</tr>
<tr>
<td>119</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
<td>4.25</td>
<td>8</td>
</tr>
<tr>
<td>121</td>
<td>1</td>
<td>4</td>
<td></td>
<td>5</td>
<td>2.5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.7</td>
<td>5</td>
</tr>
<tr>
<td>122</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4.444</td>
<td>9</td>
</tr>
<tr>
<td>123</td>
<td>4</td>
<td>5</td>
<td>2.5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3.944</td>
<td>9</td>
</tr>
<tr>
<td>124</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4.444</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n</th>
<th>19</th>
<th>18</th>
<th>18</th>
<th>20</th>
<th>16</th>
<th>14</th>
<th>18</th>
<th>17</th>
<th>17</th>
<th>16</th>
<th></th>
</tr>
</thead>
</table>

### Table B.3.2

*Exit Slip Data from Control Group*

<table>
<thead>
<tr>
<th>Student</th>
<th>Act 1</th>
<th>Act 2</th>
<th>Act 3</th>
<th>Act 4</th>
<th>Act 5</th>
<th>Act 6</th>
<th>Act 7</th>
<th>Act 8</th>
<th>Act 9</th>
<th>Act 10</th>
<th>Overall</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>201</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2.778</td>
<td>9</td>
</tr>
<tr>
<td>202</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3.143</td>
<td>7</td>
</tr>
<tr>
<td>203</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2.9</td>
<td>10</td>
</tr>
<tr>
<td>204</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>4.143</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>205</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3.714</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>206</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>207</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3.889</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>208</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2.571</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>209</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>211</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.8</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>212</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>blank</td>
<td>2</td>
<td>3</td>
<td>3.222</td>
<td>10</td>
</tr>
<tr>
<td>213</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3.375</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>214</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3.625</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>215</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>blank</td>
<td>2</td>
<td>2</td>
<td>2.556</td>
<td>10</td>
</tr>
<tr>
<td>216</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>3.1</td>
<td>10</td>
</tr>
<tr>
<td>218</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>3.6</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>219</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3.6</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3.1</td>
<td>10</td>
</tr>
<tr>
<td>221</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>222</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>223</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>224</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>blank</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2.833</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>225</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>226</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3.3</td>
<td>10</td>
</tr>
<tr>
<td>227</td>
<td>blank</td>
<td>1</td>
<td>1</td>
<td></td>
<td>2</td>
<td>3</td>
<td></td>
<td>1.75</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>228</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>229</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>3.2</td>
<td>10</td>
</tr>
<tr>
<td>230</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>231</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4.111</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

| n       | 22     | 20     | 20     | 23     | 21     | 16     | 21     | 20     | 20     | 19     |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| average | 4.273 | 4.25  | 3.9   | 2.043 | 2      | 3.563  | 2.333  | 3.05  | 3.3    | 3.263 |       |    |

*IMPACT OF USING RELEVANT CONTEXT*
Table B.3.3

Exit Slip Data from Treatment Group 2

<table>
<thead>
<tr>
<th>Student</th>
<th>Act 1</th>
<th>Act 2</th>
<th>Act 3</th>
<th>Act 4</th>
<th>Act 5</th>
<th>Act 6</th>
<th>Act 7</th>
<th>Act 8</th>
<th>Act 9</th>
<th>Act 10</th>
<th>Overall</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>301</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td></td>
<td>3.625</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>302</td>
<td>5</td>
<td></td>
<td>4</td>
<td>blank</td>
<td>4</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td>4.333</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>303</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>3.778</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>304</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
<td>3.857</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>305</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td></td>
<td></td>
<td>3.857</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>306</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>blank</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3.625</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>307</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>308</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>309</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
<td>3.571</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>310</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>311</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>312</td>
<td>5</td>
<td>5</td>
<td>blank</td>
<td>5</td>
<td></td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>313</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3.333</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>314</td>
<td>3</td>
<td>2</td>
<td></td>
<td>4</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>2.8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>315</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N/A</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>316</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>317</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>318</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td></td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
<td>4.333</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>319</td>
<td>5</td>
<td></td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>4.4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>320</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>321</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>322</td>
<td>4</td>
<td></td>
<td>4</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.25</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>323</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>4.333</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>324</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N/A</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>325</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>blank</td>
<td></td>
<td></td>
<td></td>
<td>3.833</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>14</td>
<td>16</td>
<td>8</td>
<td>10</td>
<td>14</td>
<td>4</td>
<td>17</td>
<td>9</td>
<td>14</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>3.857</td>
<td>4</td>
<td>4</td>
<td>3.8</td>
<td>3.5</td>
<td>4</td>
<td>3.824</td>
<td>3.222</td>
<td>4.429</td>
<td>3.875</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C.1

Comprehension Pretest

Quiz Instructions

Completion of this pretest is worth 10 points, no matter what grade is actually earned on it. In other words, if your score upon completion says 2/10 or 6/10, it will be changed at the end of the semester to 10/10. The only way to earn less than 10/10 is if you chose not to take the pretest or you enter answers in a way which indicates to your instructor that you didn’t even read the questions. (For instance, if you complete the entire test in less than one minute, that indicates a lack of effort.)

Please do your best to answer the questions, but do NOT get any sort of outside assistance. It is important for us to know what knowledge you are coming into the probability unit with and what knowledge you currently do not have. It is especially important that you give your best effort if you are participating in the educational research. We do not expect anyone to earn a “good” grade on this pretest, which is why we are converting everyone’s scores to a full 10 points.

Question 1

Half of all newborns are girls and half are boys. Hospital A records an average of 50 births a day. Hospital B records an average of 10 births a day. On a particular day, which hospital is more likely to record 50% or more female births?

- Hospital A (with 50 births a day)
- Hospital B (with 10 births a day)
- The two hospitals are equally likely to record such an event
- It is impossible to determine

Question 2

When two fair six-sided dice are simultaneously thrown, these are two of the possible results that could occur:

Result 1: a 5 and a 6 are obtained in any order
Result 2: a 5 is obtained on each die

- The probability of obtaining each of these results is equal
- There is a higher probability of obtaining result 1 (a 5 and a 6 in any order)
- There is a higher probability of obtaining result 2 (a 5 on each die)
- It is impossible to give an answer

Question 3

You draw one card from a standard deck of playing cards. Events A and B are defined in the following way:

Event A: Draw a King
Event B: Draw a 3

Events A and B are considered:

- Independent, but not mutually exclusive
- Dependent, but not mutually exclusive
- Both independent and mutually exclusive
- Both dependent and mutually exclusive
**Question 4**

A baker has made 260 cookies, 17% of which have chocolate chips and 8% of which have sprinkles. 35 of the 260 cookies have both chocolate chips and sprinkles. One cookie is selected at random.

(Hint: While not necessary, a Venn Diagram may help you with this problem.)

Find the probability that this cookie has either chocolate chips or sprinkles

Retype your answer: 

(Enter all answers as decimals to 3 decimal places. Do not enter fractions.)

**Question 5**

The table is based on the records of accidents compiled by a State Highway Safety and Motor Vehicles Office. The office wants to decide if people are less likely to have a fatal accident if they are wearing a seatbelt. Which of the following comparisons is most appropriate for supporting this conclusion?

<table>
<thead>
<tr>
<th>Safety Equipment in Use</th>
<th>Injury</th>
<th>ROW TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonfatal</td>
<td>Fatal</td>
</tr>
<tr>
<td>Seat Belt</td>
<td>412.368</td>
<td>510</td>
</tr>
<tr>
<td>No Seat Belt</td>
<td>162.527</td>
<td>1,601</td>
</tr>
<tr>
<td><strong>COLUMN TOTAL</strong></td>
<td>574.895</td>
<td>2,111</td>
</tr>
</tbody>
</table>

- Compare the ratio \( \frac{510}{412.368} \) and \( \frac{1,601}{164.128} \)
- Compare the ratio \( \frac{510}{577.006} \) and \( \frac{1,601}{577.006} \)
- Compare the numbers 510 and 1,601
- Compare the ratio \( \frac{2,111}{412.368} \) and \( \frac{2,111}{164.128} \)

**Question 6**

Given the probability distribution above, find the following:

\( p(10) \)
**Question 7**

Four histograms are displayed below. Match the description to the appropriate histogram.

A distribution for a set of wrist circumferences (measured in centimeters) taken from the right wrist of a random sample of newborn females is represented by:

- Histogram I
- Histogram II
- Histogram III
- Histogram IV

**Question 8**

Five faces of a fair die are painted black, and one face is painted white. The die is rolled six times. Which of the following results is more likely?

- Black side up on five of the rolls, while side up on the other roll
- Black side up on all six rolls
- Both options are equally likely
- It is impossible to tell without more information
Question 9

Suppose that the number of bug reports arriving at an IT department of a software development company follows a Poisson distribution with an average of 5 reports over a 4-day period. Use the partial probability table below to answer the following question.

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0679</td>
</tr>
<tr>
<td>1</td>
<td>0.03347</td>
</tr>
<tr>
<td>2</td>
<td>0.1255</td>
</tr>
<tr>
<td>3</td>
<td>0.0471</td>
</tr>
<tr>
<td>4</td>
<td>0.0141</td>
</tr>
<tr>
<td>5</td>
<td>0.0035</td>
</tr>
<tr>
<td>6</td>
<td>0.0008</td>
</tr>
<tr>
<td>7</td>
<td>0.0001</td>
</tr>
<tr>
<td>8</td>
<td>0.0001</td>
</tr>
<tr>
<td>9</td>
<td>0.0001</td>
</tr>
<tr>
<td>10</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Use four decimal places for your answers, since that is the number given in the table.

What is the probability that in the next four days, there will be exactly 7 bug reports?

Note: this is only a partial table. The table actually extends beyond x = 10, but it is not given to you. However, this problem can be answered with just this partial table.

Question 10

The mathematics section of the 2015 SAT was normally distributed with a mean score of 511 and a standard deviation of 120. (Enter all values for the indicated steps in your solutions)

Find the math SAT score which separates the top 25% of students from the bottom 75% of students.

The z-value is _______ and the corresponding math SAT score is _______.
### Validation of Questions for Post-Test Part 1

<table>
<thead>
<tr>
<th>Question on Post-Test</th>
<th>Question #</th>
<th>Pre-test?</th>
<th>Origin Comprehension Type</th>
<th>Concept(s)</th>
<th>Clarity/Phrasing Average Rating (n=6)</th>
<th>Address Intended Concept(s) Average Rating (n=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Half of all newborns are girls and half are boys. Hospital A records an average of 50 births a day. Hospital B records an average of 10 births a day. On a particular day, which hospital is more likely to record 80% or more female births?</td>
<td>1</td>
<td>Yes</td>
<td>SRA 14 (added choice d)</td>
<td>Reasoning Law of Large Numbers and Variability of small vs large samples</td>
<td>4.33</td>
<td>4.5</td>
</tr>
<tr>
<td>- Hospital A (with 50 births a day)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Hospital B (with 10 births a day)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- The two hospitals are equally likely to record such an event</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- It is impossible to determine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If you flip a fair coin and get heads 5 times in a row, what is the chance of getting tails on the next flip?</td>
<td>2</td>
<td>No</td>
<td>ARTIST Q0548 (edited choices)</td>
<td>Reasoning Theoretical probability and Blindness to previous outcomes for independent events</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>- Slightly greater than 50% since you are good at landing on heads</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Exactly 50%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Slightly less than 50% since tails is overdue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- 1/64 or 1.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When two fair six-sided dice are simultaneously thrown, these are two of the possible results that could occur: Result 1: a 5 and a 6 are obtained In any order Result 2: a 5 is obtained on each die</td>
<td>3</td>
<td>Yes</td>
<td>ARTIST Scale 3</td>
<td>Reasoning Probabilities of outcomes and events in a sample space</td>
<td>4.33</td>
<td>4.83</td>
</tr>
<tr>
<td>- The probability of obtaining each of these results is equal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- There is a higher probability of obtaining Result 1 (a 5 and a 6 in any order)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- There is a higher probability of obtaining Result 2 (a 5 on each die)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- It is impossible to give an answer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
You draw one card from a standard deck of playing cards. Events A and B are defined in the following way:

Event A: Draw a King
Event B: Draw a 2

Events A and B are considered:

- Independent, but not mutually exclusive
- Dependent, but not mutually exclusive
- Both independent and mutually exclusive
- Both dependent and mutually exclusive

Given two events A and B with \( P(A) = 0.7 \) and \( P(B) = 0.2 \), find \( P(A \cap B) \) if you also know that...

A.) \( P(A|B) = 0.6 \)
B.) \( P(A \cup B) = 0.85 \)

(Hint: Parts A and B use the same starting information but are completely separate questions.)

(Troubleshooting: If the symbols aren’t displaying properly, you are being asked to find A intersect B in both parts. In part B, you are told the value of A union B.)
<table>
<thead>
<tr>
<th>A. Has chocolate chips</th>
<th>Retype your answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. Does not have sprinkles</td>
<td>Retype your answer</td>
</tr>
<tr>
<td>C. Has either chocolate chips or sprinkles</td>
<td>Retype your answer</td>
</tr>
<tr>
<td>D. Has neither chocolate chips nor sprinkles</td>
<td>Retype your answer</td>
</tr>
</tbody>
</table>

(Enter all answers as decimals to 3 decimal places. Do not enter fractions.)

| 7 | No Researcher Reasoning Probability for a multistep experiment and probability of an event as the sum of probabilities of the outcomes of that event | 5 | 5 |
The table is based on the records of accidents compiled by a State Highway Safety and Motor Vehicles Office. The office wants to decide if people are less likely to have a fatal accident if they are wearing a seatbelt. Which of the following comparisons is most appropriate for supporting this conclusion?

<table>
<thead>
<tr>
<th>Safety Equipment in Use</th>
<th>Injury</th>
<th>ROW TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonfatal</td>
<td>Fatal</td>
</tr>
<tr>
<td>Seat Belt</td>
<td>412,368</td>
<td>510</td>
</tr>
<tr>
<td>No Seat Belt</td>
<td>162,527</td>
<td>1,601</td>
</tr>
<tr>
<td>COLUMN TOTAL</td>
<td>574,895</td>
<td>2,111</td>
</tr>
</tbody>
</table>

- **Reasoning**

- **Contingency Tables and Conditional Probability**

Given the table above (in the previous question), calculate the probability of a fatal accident, then determine if this probability is represented a joint probability (J), a marginal probability (M), or neither (N).

(Give your probability answer to 3 decimal places. Do not type a fraction.)

Probability = 

J, M, or N = 

Calculate P(J,J). (If your answer is in the thousands, DO NOT use commas. For instance, 12,345 should be entered 12345.)

Answer = 

Re-type your answer.
Suppose you are drawing three cards from a standard deck with replacement between draws. In order to determine the total number of outcomes for this experiment, you should use

- A permutation $P(52,3)$
- A combination $C(52,3)$
- The fundamental counting rule: $52 \times 52 \times 52$
- The fundamental counting rule: $52 \times 51 \times 50$

| 11 | No Researcher Reasoning Recognizing which counting rule is appropriate | 5 | 5 |

Overall, how well do the 11 questions on this exam covering topics in probability assess students' statistical literacy and reasoning skills? 5

Note. All ratings were made using a 5-point Likert scale from 1 (poor) to 5 (excellent).
Validation of Questions for Post-Test Part 2

<table>
<thead>
<tr>
<th>Question on Post-Test</th>
<th>Question # Pre-test?</th>
<th>Origin</th>
<th>Comprehension Type</th>
<th>Concept(s)</th>
<th>Clarity/Phrasing Average Rating (n=5)</th>
<th>Address Intended Concept(s) Average Rating (n=5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x ) ( p(x) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Given the probability distribution above, find the following:

A.) \( p(10) \)  

B.) \( p(12) \)  

One of the items on a student survey for an introductory statistics course was "Rate your intelligence on a scale of 1 to 5." Here is the distribution of this variable for the students in the class.

<table>
<thead>
<tr>
<th>Intelligence Rating (X)</th>
<th>Proportion or p(X)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>.1</td>
</tr>
<tr>
<td>3</td>
<td>.2</td>
</tr>
<tr>
<td>4</td>
<td>.3</td>
</tr>
<tr>
<td>5</td>
<td>.4</td>
</tr>
</tbody>
</table>

A.) What is the average intelligence rating for the students? (In other words, what is the expected value, \( E(X) \)?)  

B.) What is the standard deviation, \( SD(X) \)?  

Retype your answer  

Note: If needed, type your answer to three decimal places.
Suppose that you win $3 if the spinner lands on Blue, $9 if the spinner lands on Cyan, but you lose money if the spinner lands on Yellow. How much should the loss be set at in order for this to be considered a fair game?

- $3.15
- $7.00
- $6.00
- None of these

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>No</td>
<td>Researcher Reasoning Expected Value Formula and Definition of a &quot;Fair Game&quot;</td>
<td>4.6</td>
<td>5</td>
</tr>
</tbody>
</table>

Four histograms are displayed below. Match the description to the appropriate histogram.

A distribution for a set of wrist circumferences (measured in centimeters) taken from the right wrist of a random sample of newborn female infants is represented by:

- Histogram I
- Histogram II
- Histogram III
- Histogram IV

Five faces of a fair die are painted black, and one face is painted white. The die is rolled six times. Which of the following results is more likely?

- Black side up on five of the rolls; white side up on the other roll
- Black side up on all six rolls
- Both options are equally likely
- It is impossible to tell without more information

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Yes</td>
<td>CAOS 4</td>
<td>Reasoning Distribution shape for a real-world random variable</td>
<td>3.2</td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Yes</td>
<td>SRA 13 / ARTIST Q0157 (added d)</td>
<td>Reasoning Reasoning using Binomial Distribution calculation or the idea of Expected Value</td>
<td>4.2</td>
</tr>
</tbody>
</table>
A little league baseball player has a batting average of 0.470. In other words, the probability is 0.47 that the player will get a hit on his next at bat. Let X represent the number of hits out of the next 9 at bats, and suppose X follows a binomial distribution.

Find the probability that he has at least 1 hit in his next 9 at bats.

Retype your answer.

(Use 4 or more decimals at each step in your computations, but enter your answers to 3 decimal places.)

In a binomial distribution, if \( n = 50 \) and \( p = 0.62 \), find

A) The expected value, \( E(X) \)

B) The standard deviation, \( SD(X) \)

(Ensure to use 3 decimal places, if needed.)

For a Poisson distribution, an interval of length 4.5 has \( \mu = 10.8 \).

What value of \( \mu \) would you use if you changed the interval length to 6.5?

Retype your answer.

6
No
Researcher
Reasoning
Matching a verbal description with its variable,
Calculating a Binomial Probability using a formula, the meaning of "at least"

4.8
5

7a
No
Researcher
Literacy
Using the "short cut" formula for expected value of a binomial distribution

5
5

7b
No
Researcher
Literacy
Using the "short cut" formula for standard deviation of a binomial distribution

5

8
No
Researcher
Reasoning
Mu is proportional to the length of the interval for a Poisson distribution

5
4.8

9a
Yes
Researcher
Literacy
Recognizing the appropriate value of mu and x,
Reading a table to

5
5
<table>
<thead>
<tr>
<th>No</th>
<th>Researcher</th>
<th>Literacy</th>
<th>Reasoning</th>
<th>Changing mu when the interval changes, Reading a table to determine a Poisson probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>9b</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9c</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9d</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table

<table>
<thead>
<tr>
<th>( \mu )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.3679</td>
<td>0.2231</td>
<td>0.1353</td>
<td>0.0498</td>
<td>0.0183</td>
<td>0.0111</td>
<td>0.0025</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>1</td>
<td>0.3579</td>
<td>0.3573</td>
<td>0.2707</td>
<td>0.1494</td>
<td>0.0733</td>
<td>0.0500</td>
<td>0.0149</td>
<td>0.0064</td>
<td>0.0003</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>2</td>
<td>0.1839</td>
<td>0.2510</td>
<td>0.2707</td>
<td>0.2240</td>
<td>0.1465</td>
<td>0.1125</td>
<td>0.0446</td>
<td>0.0223</td>
<td>0.0015</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>3</td>
<td>0.0613</td>
<td>0.1255</td>
<td>0.1804</td>
<td>0.2240</td>
<td>0.1954</td>
<td>0.1687</td>
<td>0.0892</td>
<td>0.0521</td>
<td>0.0053</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>4</td>
<td>0.0153</td>
<td>0.0471</td>
<td>0.0902</td>
<td>0.1680</td>
<td>0.1954</td>
<td>0.1898</td>
<td>0.1339</td>
<td>0.0912</td>
<td>0.0139</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>5</td>
<td>0.0031</td>
<td>0.0141</td>
<td>0.0361</td>
<td>0.1008</td>
<td>0.1563</td>
<td>0.1708</td>
<td>0.1606</td>
<td>0.1277</td>
<td>0.0293</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>6</td>
<td>0.0005</td>
<td>0.0035</td>
<td>0.0120</td>
<td>0.0504</td>
<td>0.1042</td>
<td>0.1281</td>
<td>0.1605</td>
<td>0.1490</td>
<td>0.0513</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>7</td>
<td>0.0001</td>
<td>0.0008</td>
<td>0.0034</td>
<td>0.0216</td>
<td>0.0595</td>
<td>0.0824</td>
<td>0.1377</td>
<td>0.1490</td>
<td>0.0769</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>8</td>
<td>0.0001</td>
<td>0.0009</td>
<td>0.0031</td>
<td>0.0298</td>
<td>0.0463</td>
<td>0.1033</td>
<td>0.1304</td>
<td>0.1009</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>9</td>
<td>0.0000</td>
<td></td>
<td>0.0002</td>
<td>0.0027</td>
<td>0.0132</td>
<td>0.0232</td>
<td>0.0688</td>
<td>0.1014</td>
<td>0.1177</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>10</td>
<td>0.0000</td>
<td></td>
<td>0.0000</td>
<td>0.0008</td>
<td>0.0053</td>
<td>0.0104</td>
<td>0.0413</td>
<td>0.0710</td>
<td>0.1236</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Use **FOUR decimal places** for your answers, since that is the number given in the table!

A.) In the next four days, there will be exactly 7 bug reports.

B.) In the next four days, there will be at least 2 bug reports.

C.) Tomorrow (1 day), there will be no bug reports.

D.) Next week (7 days), there will be less than 4 bug reports.

**Note:** This is only a partial table. The table actually extends beyond \( x = 10 \), but it is not given to you. However, all problems can be answered with just this partial table.
Suppose $X$ is a uniformly distributed variable with values between 35 and 75.

(Enter your answers correct to 3 decimal places, if needed. Do not use fractions.)

(Hint: Sketching a picture will likely help you answer these questions.)

A. Find $P(X > 48)$ Retype your answer. 

B. Find $P(52 < X < 67)$ Retype your answer. 

C. Find $P(68 \leq X \leq 80)$ Retype your answer. 

D. Find $E(X)$ (also called $\mu$) Retype your answer. 

An infant's height at his one-year check-up provided a standardized ($z$) score that was -0.57. What does this score tell about how tall the infant was in relation to all other one year olds?

- He was in the bottom 57% of all heights of one year olds.
- He was 57% smaller than an average one year old.
- He was slightly more than half a standard deviation below the average height of a one year old.
- $1 - .57 = .43$, so 43% of one year olds are smaller than him.

Find the following probability for the standard normal distribution.

Note: Use the same number of decimal places as are on your table and enter all steps for your computation.

$P(-1.29 < z < 0.46) =$ 

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning</td>
<td>Calculating probability using a Uniform Distribution</td>
</tr>
<tr>
<td>No</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning</td>
<td>Calculating probability using a Uniform Distribution when some values of $X$ are unreasonable</td>
</tr>
<tr>
<td>No</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning</td>
<td>Using the Expected Value formula for a Uniform Distribution</td>
</tr>
<tr>
<td>No</td>
<td>4.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning</td>
<td>Understanding what a z-score represents</td>
</tr>
<tr>
<td>No</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning</td>
<td>Reading the standard normal table, Calculating a probability using the normal distribution</td>
</tr>
<tr>
<td>No</td>
<td>4.8</td>
</tr>
</tbody>
</table>

| Researcher | Literacy | 5 |
Find the following values of $a$ using the standard normal table:
(Note: Use an appropriate number of decimal places, based on the values you see in the table.)

A.) $P(Z > a) = .352 \Rightarrow a = \phantom{0000}$
Retype your answer.

B.) $P(-a < Z < a) = .99 \Rightarrow \pm a = \pm \phantom{0000}$
Retype your answer.

The mathematics section of the 2015 SAT was normally distributed with a mean score of 511 and a standard deviation of 120. (Enter all values for the indicated steps in your solutions)

A.) Find the probability that a randomly selected student who took the exam had a math SAT score less than 450.

$P(X < 450) = P(Z < \phantom{0000}) = \phantom{0000}$

B.) Find the math SAT score which separates the top 25% of students from the bottom 75% of students.

The $z$-value is $\phantom{0000}$ and the corresponding math SAT score is $\phantom{0000}$

Which of the following describes an exponential random variable?

- The number of people out of 20 who reacted to their COVID vaccination with a fever
- The number of people who are vaccinated at a walk-in clinic during an 8-hour shift
- The time between patient arrivals at a walk-in vaccination clinic
- The height of females who receive their COVID vaccine at a walk-in clinic

Using the complement rule, Reading the standard normal table in reverse, Determining a normal probability

13b

No
Researcher
Literacy
Using symmetry and the complement rule, Reading the standard normal table in reverse

14a
Yes
Researcher
Reasoning
Converting to a z-score, Determining a normal probability

14b

Yes
Researcher
Reasoning
Reading the standard normal table in reverse, Determining an observation based on a z-score

15

No
Researcher
Literacy
An exponential random variable measures the length between occurrences or the length until the first occurrence

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 13b | No | Researcher
|     | Literacy | Using symmetry and the complement rule, Reading the standard normal table in reverse |
|     | 5    |   |
| 14a | Yes | Researcher
|     | Reasoning | Converting to a z-score, Determining a normal probability |
|     | 5    | 5  |
| 14b | Yes | Researcher
|     | Reasoning | Reading the standard normal table in reverse, Determining an observation based on a z-score |
|     | 5    |   |
| 15  | No  | Researcher
|     | Literacy | An exponential random variable measures the length between occurrences or the length until the first occurrence |
|     | 4.8 | 4.8 |
For an *exponentially* distributed variable, X, with $\mu = 7$, the calculation $1 - e^{-2T}$ would be the answer to which of the following questions:

<table>
<thead>
<tr>
<th>Question</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 No Researcher Literacy How the exponential formula relates to probability</td>
<td>5</td>
</tr>
<tr>
<td>17 No Researcher Reasoning The probability at one point is always 0 for continuous random variables, Recognizing which distributions are continuous and which are discrete</td>
<td>5</td>
</tr>
</tbody>
</table>

Overall, how well do the 17 questions on this exam covering probability distributions assess students' statistical literacy and reasoning skills?

4.6

*Note.* All ratings were made using a 5-point Likert scale from 1 (*poor*) to 5 (*excellent*).
APPENDIX C.3

SATS-36 Explanation

The pre- and post-versions of the Survey of Attitudes Toward Statistics (SATS-36) are copyrighted instruments and may be obtained with the direct permission of Candace Schau by completing an online form found at https://www.evaluationandstatistics.com/register.
APPENDIX C.4

Qualitative Script / Interview Questions

Introduction
Hello, my name is xxxxxxxxxxx. I’ll be leading our [interview/focus group] today. As I mentioned over email and as you might recall, I am part of the team doing a research study about the impact of relevant contextual examples on student comprehension and attitude.

I would like to thank you for your time today and being willing to speak with us about the probability unit in your statistics course. Your feedback will be used for our doctorate dissertation. Understand that there are no wrong answers, nor do we have any expectations on how you might answer. Your honest response is essential. You will not hurt our feelings nor get anyone in trouble.

We are planning to keep this [interview/focus group] to [time allotment] minutes. Let us know if you need a break or need to stop for any reason during that time. Does this work for you?

With your permission, I’d like to record this [interview/focus group]. The recording will only be used to help us in our research, and it won’t be shared with anyone except the three of us and our professors if necessary. Recording this call helps us ensure that we capture the information accurately.

Finally, I want to confirm that you’ve received a participant consent document — is that correct? (Note: Have a copy of the consent for reference.)

Great. Do you have any questions for me before we continue?

Comprehension
Now, I’d like to ask you a few questions about your comprehension of the probability concepts in your statistics course.

1. Do you feel as if you have a strong, moderate, or weak comprehension of probability topics after these past few weeks?
2. How would you describe your change in comprehension over the probability unit?
3. Which element of the instructional process (video lectures, in-class activities, peer support, instructor feedback) was most crucial to your comprehension of probability topics? If any.
4. How did the context used during the in-class activities affect how you comprehended the concepts? And it is ok if it did not have any affect. *(Note: Have an example of a probability activity using relevant context vs standard context.)*

5. Were there any outside circumstances that you feel affected your exam performance? For example: sickness; death in the family; recent break-up; unexpected call into work which affected study time; etc.

**Attitude**

I am going to shift the questions slightly. I will now be asking you questions associated with your attitude toward statistics.

6. Do you have an overall positive, neutral, or negative attitude toward statistics?

7. How would you describe your change in attitude over the probability unit?

8. Which element of the instructional process (video lectures, in-class activities, peer support, instructor feedback) was most crucial to your attitude toward statistics?

9. How did the context used during the in-class activities affect your attitude toward statistics? *(Note: Have a list of interest topics used to develop the relevant context.)*

10. Were there any outside circumstances that you feel affected your attitude toward statistics? For example: likeability/credibility of the instructor; influence of friends/family; etc.

I want to thank you for your participation so far. We are almost done. One last question.

11. Do you have anything more you would like to share with us about how the use of relevant context affected your change in comprehension of the probability unit or attitude toward statistics?

On behalf of we would like to thank you for your time. We appreciate you sharing your experiences and thoughts with us. They will be valuable as we add to the collective knowledge about the use of relevant contextual examples in collegiate-level courses. If you have questions about this research study, please feel free to contact us.
APPENDIX D.1

Treatment Group Activities

Activity 1:

**In-Class Activity: Law of Large Numbers**

**Egg Russian Roulette with Jimmy Fallon**

Watch this video of Jimmy Fallon and Tom Cruise:
https://www.youtube.com/watch?v=y6jb0cOYqAI

**Part I: Questions to think about with your group...**

How likely is it that Tom loses that quickly?

Jimmy says, “That never happens” when Tom gets a raw egg on the first attempt. What is the probability Tom gets a raw egg on his first attempt?

Given that the first egg was raw, how likely is it that Jimmy was safe with the next egg?

How does the probability of getting egg on your head change each time an egg is selected?

Do you think each player has the same probability of winning the game when guest of the show goes first, or does one player have an advantage over the other?

As a class, let’s vote...Is it better to go first, go second, or does it not matter?
Part II: The Yolk’s on You!
With your partner, play the egg roulette game 5 times. Record in the table which player won and the round at which the other player lost. (See example) Guest of the show must take the first turn in all 5 games! (To avoid bias, take a new carton of eggs for each game and reset the game for the next team when you are done. Don’t use the same carton twice!)

<table>
<thead>
<tr>
<th></th>
<th>Player 1: (Guest)</th>
<th>Player 2: (Jimmy Fallon)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example Tom Cruise</td>
<td>Loses in round 3</td>
<td>WINNER</td>
</tr>
<tr>
<td>Game 1 (Set ____ )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game 2 (Set ____ )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game 3 (Set ____ )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game 4 (Set ____ )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game 5 (Set ____ )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What do you notice about the round number in which the guest loses? What about the round number in which Jimmy loses?

There are 12 eggs, but will the game ever go to round 12? Why or why not?

Based on your trials, what is the probability of Jimmy winning? Express your answer as a decimal.

Based on your trials, do you now believe it is better to go first, second, or does it not matter?

Based on your trials, what is the probability of losing in exactly 3 rounds? What is the probability of losing in 9 rounds or more?

Part III: Team up with another group, combine your results and answer these last 3 questions (starting with the words, “Based on your trials”) again… Have your answers changed? Why or why not?
**Part IV:** Enter the data in your table (from Part II) in the Excel Spreadsheet.

Look at the class results for overall wins:

<table>
<thead>
<tr>
<th>Wins for Guest:</th>
<th>Wins for Jimmy:</th>
</tr>
</thead>
</table>

Based on our class data, what is the probability of Jimmy Fallon winning egg roulette? Is this a theoretical or experimental probability? How do you know?

What happened to the probability as we used more trials?

Do you think we have enough data to predict the theoretical probability of Jimmy Fallon winning? Why or why not? If not, how many more trials would you recommend?

Now let’s officially answer the question...Is it better to go first, go second, or does it not matter?

Is it more likely for Jimmy to win 80% of the time if he played only 5 games or if he played 100 games? Explain.
Let’s take a look at the class results for rounds in which the game ends:

What is the most likely number of rounds for a game to end? What is the least likely?

Theoretically, just over ¼ of the games played should take 9 or more rounds.

- Scenario A: 5 games were played
- Scenario B: 50 games were played
- Scenario C: 500 games were played

In which scenario are you more likely to see that 60% of the games played took 9 or more rounds?

In which scenario are you more likely to see that 25% of the games played took 9 or more rounds?

How does the Law of Large Numbers relate to your answers above?
Activity 2

In-Class Activity: OR Probability
Travel

Part 1: Without looking anything up online, give your best guess on the two following trivia questions related to the world’s most popular tourist destinations.

1) In 2019, the following 5 countries had the highest international tourism, based on number of tourist arrivals. Which country out of the top 5 do you think ranked #1?
   a. China
   b. France
   c. Italy
   d. Spain
   e. United States

2) In 2019, the following 5 attractions were the most visited in the world. Which attraction out of the top 5 do you think had the greatest number of visitors?
   a. Colosseum (Rome, Italy)
   b. Eiffel Tower (Paris, France)
   c. Louvre Museum (Paris, France)
   d. Statue of Liberty (New York City, USA)
   e. Vatican Museums (Rome, Italy)

As a class, we are going to vote on what we think the correct answer is. Don’t be influenced by what other people are choosing – stick with your initial answers! Keep track of class votes in the tables below.

<table>
<thead>
<tr>
<th>Country Selected</th>
<th>China</th>
<th>France</th>
<th>Italy</th>
<th>Spain</th>
<th>United States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Votes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attraction Selected</th>
<th>Colosseum</th>
<th>Eiffel Tower</th>
<th>Louvre</th>
<th>Statue of Liberty</th>
<th>Vatican Museums</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Votes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Part 2: After we reveal the correct answers to the trivia questions above, use our class data to answer the questions that follow.

3) Let the event $C$ represent getting the correct answer to the country question (#1). If we were to select a student at random from our class, what is $P(C)$?

4) Let the event $A$ represent getting the correct answer to the attraction question (#2). If we were to select a student at random from our class, what is $P(A)$?

5) Suppose that we selected a student at random, and wanted to find $P(C \cup A)$. Explain in words what this probability would mean.

6) Unfortunately, we cannot just add together $P(C)$ and $P(A)$ to get $P(C \cup A)$. This is because the events $C$ and $A$ are not mutually exclusive! Explain why the events $C$ and $A$ are not mutually exclusive.

7) With “OR” probability, it can be really helpful to draw out the problem using a Venn Diagram. Fill out each blank in the picture below with the correct number of votes from our class.

8) Calculate $P(C \cup A)$, using either your Venn Diagram above, or the formula shown below.
$$P(C \cup A) = P(C) + P(A) - P(C \cap A)$$

9) The complement of an event represents everything not included within the original event. What is the complement of your answer to #8? Explain what this probability represents, in terms of the trivia voting from earlier.
Part 3: Circle your top 10 dream destinations from the list shown below. Then, compare selections with a partner to answer the questions that follow.

The travel website Planetware put out a 2020 “Bucket List Destinations” with the top 20 places to visit in your lifetime. You can pull up the site to see pictures if you’d like, but then circle the 10 places you’d most like to visit, if money and time were not an issue.

<table>
<thead>
<tr>
<th>Machu Picchu, Peru</th>
<th>Pyramids of Giza, Egypt</th>
<th>African Safari, Kenya</th>
<th>Taj Mahal, India</th>
<th>Paris, France</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon Rainforest, Brazil</td>
<td>Santorini, Greece</td>
<td>Bora Bora, French Polynesia</td>
<td>Northern Lights, Iceland</td>
<td>New York City, USA</td>
</tr>
<tr>
<td>Venice, Italy</td>
<td>Angkor Wat, Cambodia</td>
<td>Niagara Falls, Canada</td>
<td>Grand Canyon, USA</td>
<td>Great Barrier Reef, Australia</td>
</tr>
<tr>
<td>Tahiti, French Polynesia</td>
<td>Iguazu Falls, Argentina</td>
<td>Rome, Italy</td>
<td>Dalmatian Coast, Croatia</td>
<td>Cinque Terre, Italy</td>
</tr>
</tbody>
</table>

Let $P(Y)$ represent the probability that a site above is on your list, and let $P(Z)$ represent the probability that a site is on your partner’s list. Obviously, these probabilities are both equal to $10/20$, since you were instructed to select 10 sites.

10) Compare and discuss your selections with a partner. Place check marks on the table above to represent places that you both selected. Then, find $P(Y \cap Z)$.

11) Using your answer to #10, calculate $P(Y \cup Z)$.

12) Find $P(Y^C \cap Z^C)$. Explain what this probability represents in the context of the bucket list.
Part 4: Use the following table to get additional practice with “OR” probability.

In 2019, Big 7 Travel compiled a list of users’ top 50 bucket list destinations. Let the event \( B \) = the destination features beaches, and the event \( E \) = the destination is located in Europe. Results are compiled in the table below.

<table>
<thead>
<tr>
<th></th>
<th>Beaches Featured</th>
<th>Beaches Not Featured</th>
</tr>
</thead>
<tbody>
<tr>
<td>European Location</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Non-European Location</td>
<td>10</td>
<td>26</td>
</tr>
</tbody>
</table>

13) Find \( P(B \cap E) \)

14) Find \( P(B \cup E) \)

15) Find \( P(B \cup E^c) \)
Activity 3

In-Class Activity: Tree Diagrams
Birth Control

Part 1: Carefully read the information below to familiarize yourself with today’s lesson context.

Many different forms of contraception exist to prevent an unintended pregnancy. Effectiveness of a contraceptive is typically given as a rate over a typical year of use. For example, a method that is 95% effective means that over the course of one year, approximately 5 out of 100 women would still become pregnant while using that method of birth control.

Birth control also includes two different rates of effectiveness. “Perfect Use” refers to how well the method works if the individual uses it correctly every single time. “Typical Use” refers to how well the method works while accounting for human error. For example, someone may occasionally forget to take a daily birth control pill, or may be on an antibiotic, which reduces how well the pill works. Someone using a condom may not put it on correctly, or may wait too long before putting it on.

Other methods of birth control, such as an Intrauterine Device (IUD), do not contain a significant risk of human error. For this activity, we will focus on the “Typical Use” effectiveness rate of two common forms of birth control.

- Condoms have a “Perfect Use” effectiveness rate of approximately 98%, but a “Typical Use” effectiveness rate of only 82% (p = 0.82).
- Daily birth control pills have a “Perfect Use” effectiveness rate of more than 99%, but a “Typical Use” effectiveness rate of only 91% (p = 0.91).

Part 2: Use the probability rates given above to answer the following questions.

1) Suppose that a woman uses condoms as her only form of birth control. Assuming the “Typical Use” rate given above, what is the probability she experiences an unintended pregnancy within a year?
2) Below is a tree diagram showing two women who use condoms as their only form of birth control. Over the course of 1 year, what is the probability that at least 1 of the women will experience an unintended pregnancy? Show two ways to arrive at this answer.

3) Expand the tree diagram to include 3 women who use condoms as their only form of birth control. Over the course of 1 year, what is the probability that at least 1 of these 3 women will experience an unintended pregnancy?

4) Imagine expanding this tree diagram to include 20 women who use condoms as their only form of birth control. Over the course of 1 year, what is the probability that at least 1 of these 20 women will experience an unintended pregnancy?

5) Suppose that a woman uses the birth control pill as her only form of contraception. Assuming the “Typical Use” rate given above, what is the probability she experiences an unintended pregnancy within a year?

6) Suppose that 20 women use the birth control pill as their only form of birth control. Over the course of 1 year, what is the probability that at least 1 of the women will experience an unintended pregnancy?
Part 3: By combining 2 different forms of birth control, you can drastically reduce the chances of an unintended pregnancy. Answer the following questions, using a tree diagram to assist in calculations.

7) The tree diagram below represents the possible outcomes for a woman on a birth control pill, while also using condoms regularly. Fill out the tree diagram, using the “Typical Use” probabilities from earlier. Then, calculate the probability for each of the 4 outcomes.

8) Based on your tree diagram above, what is the probability that a woman with “Typical Use” of both forms of contraception will experience an unintended pregnancy within a year? Show your work.

9) Keep in mind that this probability can be reduced even further by more careful use of the chosen birth control method. Based on the probability you calculated in #8, how many times lower is the risk of unintended pregnancy when using both forms of birth control, when compared to using condoms alone?
Part 4: Let’s think back to the Egg Roulette example we did last week.

10) Create a tree diagram for three rounds of Egg Roulette, and determine the likelihood that Tom Cruise loses in just three rounds.

11) If you were to add a fourth round of Egg Roulette, would new branches be drawn at the end of every path in your diagram above? Why or why not? How many total outcomes would there be at the end of the 4th round?
Activity 4

In-Class Activity: Conditional Probability
Pets and Attraction

Part 1: Read this excerpt from “The Cute Dog Effect on Sex, Money, and Justice”.
Then answer the questions about the article.

Cute Dogs Make Men Sexier
Antoine, a handsome young Frenchman in his early twenties, approaches an attractive young woman on the street. A cute energetic black and white dog named Gwendu is with him.
"Hello," he says to the woman. "My name's Antoine. I just want to say that I think you're really pretty. I have to go to work this afternoon, but I was wondering if you would give me your phone number? I'll phone you later and we can have a drink together someplace." She hesitates for a second, looks at him and then his dog. "Oui," she says and pulls a pen from her purse.
The truth is that Antoine is actually a confederate in an experiment designed by Serge Ciccotti and Nicolas Guéguen of the Université de Bretagne-Sud. (As described by Psych Today blogger Gaad Saad, Guéguen is known for discovering female hitchhikers with large breasts are more likely to get picked up by male drivers). Now he is studying whether dogs make men sexier. Over several weeks, Antoine chatted up 240 randomly selected young women. On half of these approaches, he was alone and on the other half he was accompanied by Gwendu.
Did the cute dog increase Antoine's sexual charisma? *Mais oui.* While about ten percent of the women gave him their phone numbers when he was by himself, nearly thirty percent of them fell for Antoine's line when he was accompanied by *le chien.* Poof....the cute dog instantly tripled Antoine's sex appeal.

1) For the rest of this section, let’s use A to denote the event that Antoine was *alone* when approaching a woman, D to denote the event that Antoine had his *dog* when approaching a woman, and N to denote the event that he received a woman’s phone *number.* What statistics or probabilities did you see in the excerpt and what symbolic notation or formula could you use with them?
2) Fill in the probability table below. You may need to do a few calculations to complete it.  
(Hint: Do the 10% and 30% have a place on the table, or are they describing something else?)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Nc</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3) Use the information in the table to answer the following questions:
A.) If one of these 240 women were randomly chosen, what is the probability that she was approached by Antoine alone and that she did not give him her phone number?

B.) Is this answer to Question A the same as \( P(A) \times P(N^c) \)? Why or why not?

C.) Are the events of being alone and not receiving a phone number independent, dependent, and/or mutually exclusive? Explain your choice(s).

D.) Out of the 240 women Antoine approached, how many of them actually gave him their phone number?

E.) Given that Antoine received a phone number, what is the probability that he had the dog with him?
Part 2: Psychology Today published a related article titled, “Does Having a Pet Make You More Attractive?”. [https://www.psychologytoday.com/us/blog/animals-and-us/201511/does-having-pet-make-you-more-attractive](https://www.psychologytoday.com/us/blog/animals-and-us/201511/does-having-pet-make-you-more-attractive) This article detailed a collaborative study between Petsmart and Match.com in which 1,210 participants took an online survey to determine how pets entered into their dating lives. 60% of the participants were women (W) and 40% of the participants were men (M). All participants owned at least one pet. Other results from the study were:

- 35% of women and 26% of men said they were more attracted to someone who owned a pet (O).
- 50% of women and 25% of men judged their dates based on how that person responded to their pet (R).
- 76% of women and 60% of men evaluated dates based on whether their own pets like the person (L).
- 64% of women and 49% of men said they were more attracted to a person if they owned a rescue animal (A).
- 75% of women and 54% of men said they would not date someone who didn’t like pets. (D)

Select one of these bullet points and create a probability table based on the data.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>

Enter your letter indicated in your chosen bullet point in the questions below, determine the probabilities, state if they are joint, marginal, or neither, and then give a verbal description of what the probability indicates:

\[ P(\_) \]

\[ P(W \cap \_) \]

\[ P(M \cup \_^c) \]

\[ P(W|\_) \]
Batch testing is a technique where samples from multiple sources are mixed and analyzed together. Medical labs will often use batch-testing on bloodwork, urine samples, or other labwork to deliver fast results at a reduced price. During summer 2020, the FDA authorized batch testing for COVID, when at-home rapid tests were not yet available and hospitals were running out of COVID tests due to the unexpected number of patients.

**Part 1: Batch Testing on Bloodwork**

Suppose that 12 people need to be given a blood test to determine whether they have a certain disease. Assume that each person has a 10% chance of having the disease, independent of the results of any other person. Consider two different methods for conducting the tests:

- **Method A:** Give an individual blood test to each person.
- **Method B:** Combine blood samples from all 12 people into one batch and test the batch.
  - If at least one person in the 12 has the disease, then the batch test result will be positive, and then all 12 people will need to be tested individually.
  - If nobody has the disease, then the batch test result will be negative, and no additional tests will be needed.

a. Let the random variable $X$ represent the total number of tests needed with Method B (batch testing). Determine the probability distribution of $X$.

<table>
<thead>
<tr>
<th>Number of Tests (X)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. Is the variable $X$ discrete or continuous? Explain.

c. If you implement Method B once, what is the probability that the number of tests needed will be smaller than it would be with Method A?
d. Determine the expected value of $X$.

e. Interpret this expected value in the context of the problem.

f. If you had to test thousands of groups of 12 people, which method - A or B - is better? Why?

Now consider Method C: We will first randomly divide the 12 people into two groups of 6 people. Within each group, we will combine blood samples from the 6 people into one batch. Then, we test both batches separately.

- As before, a batch will test positive only if at least one person in the group has the disease. Any batch that tests positive requires individual testing for the 6 people in that group.
- As before, a batch will test negative if nobody in the group has the disease. Any batch that tests negative requires no additional testing.

g. Let the random variable $Y$ represent the total number of tests needed for a batch of 12 with Method C (12 people separated into batches of 6). Determine the probability distribution of $Y$.

<table>
<thead>
<tr>
<th>Number of Tests (Y)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

h. If you implement Method C once, what is the probability that it will require fewer tests than Method A?

i. Determine the expected value of $Y$. Is Method C better than Method B? Why or why not?

j. Calculate and interpret the standard deviation of $Y$. 
Part 2: Batch Testing with COVID

Discuss the following questions with your group:

1.) When the FDA authorized batch-testing for COVID, one of the guidelines was that in order to create a pooled sample, each individual must be asymptomatic. Why do you think they set this guideline?

2.) Because there is asymptomatic spread of COVID, a second guideline set forth by the FDA was that the positivity rate in the local community should be below a certain threshold. Why do you suppose this was one of their guidelines? (The positivity rate is the number of people who test positive divided by the total number of people being tested.)

3.) A “false negative” result occurs when a patient has a disease but the test indicates that they do not. Batch testing has a higher false negative rate when more samples are combined together, because the amount of detectable disease in the overall sample will appear more diluted than in each individual sample. For COVID, in order to prevent too many false negative results, batch testing was limited to 5 samples being pooled together. Currently, in St. Louis, the positivity rate is ____%. Assuming a laboratory follows the FDA guidelines of mixing 5 samples in a batch, what is the probability that a batch test will need to be redone with individual tests?

4.) A St. Louis employer wants to test 5 of their employees before returning in-person next week. If each COVID test costs $10, would it be cheaper to do batch testing or cheaper to just test everyone? Would your answer change if...
   a.) The positivity rate was lower
   b.) The number of samples in a batch was higher
   c.) The cost of the test was cheaper
Activity 6

In-Class Activity: Probabilities and the Binomial Distribution

Part 1 (to be completed before coming to class): Let’s play the game 2 Truths and a Lie. In the traditional version of this game, one person states three facts, two of which are true and one is a lie. The remaining players try to guess which one is the lie.

We are going to put a fun statistical twist on this game. Before class, you’ll do some research and find verbal descriptions of 3 unlikely events. Two of them should have probabilities that are more common than 1 in a million (.000001 or more). The other event is actually going to be rarer than 1 in a million (less than .000001). Try to pick events whose probabilities surprised you and that most people won’t already know. Pick something that’s fun and interesting to you. Try to avoid using the same events as other people in the class, if possible. The more variety we have as a class, the more fun this will be. Note the sources of where you got your data from, but keep both the sources and the probabilities a secret from everyone for now. (We'll ask you for them later.) Write the three events below next to the letters A, B, and C in any order so we don't know where the "lie" is.

Example of three unlikely events:

A. The chance of winning an Olympic Gold medal in your lifetime
B. The chance of being killed by a meteorite
C. The chance of becoming a movie star

A:______________________________________________________________________
B:______________________________________________________________________
C:______________________________________________________________________

Turn in this sheet to your instructor as you walk into class next time, and make sure you bring the likelihoods of all three events on a separate sheet of scratch paper.
In-Class Activity: Probabilities and the Binomial Distribution

Part 2: Hand in your worksheet and tell your instructor which letter was the “lie” (i.e., the rarest probability) on the master solutions spreadsheet.

Part 3: Walk around the classroom and look at 5 or more worksheets that your peers have submitted, and make your best guess as to which was the “lie”. But here are the rules: No researching online...Just use your intuition. It's perfectly okay if you guess wrong. In fact, you should be wrong about 2/3 of the time if you really are just guessing! Also, don't let other students' guesses influence yours. Everyone is guessing on these. Record the letter of your guesses in the appropriate slot below.

Set 1:   Set 2:   Set 3:   Set 4:   Set 5:   Set 6:
Set 7:   Set 8:   Set 9:   Set 10:  Set 11:  Set 12:
Set 13:  Set 14:  Set 15:  Set 16:  Set 17:  Set 18:
Set 25:  Set 26:  Set 27:  Set 28:  Set 29:  Set 30:
Set 31:  Set 32:  Set 33:  Set 34:  Set 35:  Set 36:
Set 37:  Set 38:  Set 39:  Set 40:  Set 41:  Set 42:
Set 43:  Set 44:  Set 45:  Set 46:  Set 47:  Set 48:

Part 4: When your instructor makes the master solutions spreadsheet available, “grade” your work and fill in the blanks below:

I guessed correctly on _____ out of _____ sets.
**Part 5:** Join with your group and fill in the table below:

<table>
<thead>
<tr>
<th>Group Member</th>
<th># of Correct Guesses</th>
<th>Total # of Guesses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What was the theoretical probability of one correct guess?

What was the experimental probability of one correct guess?

**Part 6:** Let $X$ be a random variable that represents the number of correct guesses out of all the attempts for members in your group. What type of distribution does $X$ have? Justify your response, indicating values for any of the variables that describe the distribution. (If needed, use the value from the theoretical probability rather than the experimental probability.)

**Part 7:** Using the distribution identified above, answer the following questions:

1.) What is the probability that your group would have guessed them all correctly?

2.) What is the probability that another group, guessing the same amount as you, would have more correct responses than your group? (Hint: You may want to use technology to help you here.)

3.) What is the expected number of correct guesses, give or take how much?

4.) Is guessing this expected number actually a realistic possibility? Why or why not?
Activity 7

In-Class Activity: Poisson Distribution
Conspiracy Theories

Part 1: Read the introduction below. Then, read the instructions for the group activity.

In 2016, Dr. David Grimes of Oxford University published a mathematical model that can be used to help disprove conspiracy theories. His model states that the more people who are involved on the inside of a conspiracy, and the more time that passes, the greater the probability that a leak will occur, and reveal the truth to the public. Dr. Grimes focused on 4 common conspiracy theories, one of which was the theory that the moon landing was a hoax. Based on the number of people that would need to be involved and keep quiet, Dr. Grimes predicted that the moon landing would have been discovered as a hoax in less than 4 years if it really was faked. The math that he used to arrive at this conclusion is based on the Poisson Distribution. We will be examining an oversimplified version of his formula in today’s activity!

Group Activity Instructions

● Suppose all members of your group are on the “inside” of a conspiracy, and you’ve been asked to hold a press-conference.
● No “insider” on the conspiracy is allowed to speak words that start with the letter “T”. Every time they do, it represents a “leak” that could lead to the conspiracy getting out into the public.
● For 4 minutes, group members will go around in a circle, describing something good that they recently had to eat. Go into excessive detail – when was it? Where were you? What made it so good? While each person is talking, the others should closely listen and keep track of how many times the speaker used words that start with “T”. You should aim to talk for a full minute each.

After the Activity...

● Once everyone has had a turn, let each person in your group know how many times they may have had a leak in their 1 minute of talking.
● Take the average number of leaks for your group. This value will serve as your group’s expected number of leaks per minute, or \( \mu \).

My Group’s \( \mu \): _______
Suppose that the probability of a leak was the same over any equal interval of time. This situation can be modeled by a Poisson Distribution! Use your group’s \( \mu \) to calculate the probabilities that follow.

**Part 2: Answer the following questions using your group’s \( \mu \)**

1) In a 1-minute period, what is the probability that exactly 6 leaks will occur?

2) In a 1-minute period, what is the probability that more than 2 leaks will occur?

3) In a 4-minute period, what is the probability that exactly 30 leaks will occur? Why is this number so small?

**Part 3: Answer the following questions using the information provided below.**

The challenge in Dr. Grimes’ study was calculating \( \mu \), or the expected number of leaks in a 1 year period. To estimate this value, Dr. Grimes used data from historic conspiracy theories that actually ended up being true, such as the NSA PRISM Project, which was revealed by Edward Snowden. He examined the number of people involved in each conspiracy, and how long it took before the conspiracy was leaked to the public.

In the case of the moon landing being a hoax, supposing that only a tight circle of NASA and government employees knew the truth, we can obtain \( \mu = 0.10 \). This would mean that we can expect an average of 0.10 leaks over a 1-year period.

4) In a 1-year period, what is the probability that there will be no leaks? Use the Poisson Formula.

5) In a 1-year period, what is the probability that there will be exactly 3 leaks? Use the Poisson Formula.
6) In a 1-year period, what is the probability that there will be either 1 or 2 leaks?
Use the Poisson Table Provided in Question 9.

7) Over a 30-year period, what is the probability that there will not be any leaks?
Use the Poisson Formula.

8) Over a 30-year period, what is the probability that there will be exactly 4 leaks?
Use the Poisson Formula.

9) Over a 30-year period, what is the probability that there will be at least 2 leaks?
Use the Poisson Table provided below.

| \( \lambda \) |
|---|---|---|---|---|---|---|---|---|
| 0.1 | 1.0 | 2.0 | 3.0 | 4.0 | 5.0 | 5.3 | 10.0 | 20.0 |
| 0 | 0.9048 | 0.8679 | 0.8353 | 0.8048 | 0.7763 | 0.7500 | 0.7253 | 0.7023 |
| 1 | 0.0905 | 0.0867 | 0.0835 | 0.0805 | 0.0776 | 0.0753 | 0.0725 | 0.0702 |
| 2 | 0.0045 | 0.0100 | 0.0160 | 0.0220 | 0.0280 | 0.0340 | 0.0400 | 0.0460 |
| 3 | 0.0002 | 0.0006 | 0.0011 | 0.0016 | 0.0021 | 0.0026 | 0.0031 | 0.0037 |
| 4 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 5 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 6 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 7 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 8 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 9 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 10 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 11 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 12 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 13 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 14 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 15 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

10) It has been approximately 53 years since the moon landing. What is the probability that someone would have leaked the truth in the past 53 years? So what do you think...Was the moon landing real or faked?
Part 4 (optional, if time): Another common conspiracy involves climate change as a man-made issue versus a natural occurrence on earth. In 2016, Dr. Grimes described the problem in this way:

**Climate change conspiracy**—Climate-change denial has a deep political dimension. Despite the overwhelming strength of evidence supporting the scientific consensus of anthropogenic global warming, there are many who reject this consensus. Of these, many claim that climate-change is a hoax staged by scientists and environmentalists, ostensibly to yield research income. Such beliefs are utterly negated by the sheer wealth of evidence against such a proposition, but remain popular due to an often-skewed false balance present in partisan media, resulting in public confusion and inertia.

Scientists have been studying climate change and began publishing papers about fossil-fuel industry effects since the 1950s. However, this did not receive mainstream attention until much later. According to Dr. Grimes’ research, there are approximately the same number of scientists currently working on climate change as there were scientists working for NASA in the 1960’s, and therefore using the same estimate for \( \mu \) seems reasonable. Again, assume \( \mu = 0.10 \). This would mean that we can expect an average of 0.10 leaks over a 1-year period.

11.) Come to a consensus in your group of an estimate (number of years) for how long man-made climate change denial has been occurring. (This is just a best guess. Each group will likely think differently.) Then convert your value of \( \mu \) for an interval of this length.

12.) Assuming that the climate scientists are wrong and that it’s been a scientific cover-up for all these years, what is the likelihood that there would have been at least one conspiracy-breaking leak by now? (You’ll need to use the Poisson formula for this!)

13.) Discuss the practical implications of this calculation with your group. Do you believe in the calculation or do you believe in the cover-up? Did our simplification of \( \mu \) affect the results enough so you don’t know what to believe?
Activity 8

In-Class Activity: The Uniform Distribution
The Price is Right

Part 1: Guess that price!

1. Air Jordans are often sold at a price point of over $400. What is the actual cost Nike incurs to produce the sneaker? The cost falls somewhere in the range between $5 and $125. Can you guess the correct price, give or take $10?

My single number best guess is _______.

The range in which I would have a good guess goes from _____ to _____.

My guess was _______________ (good / bad)

2. A TI-84 graphing calculator is sold for around $120. Can you guess the actual cost to produce the calculator, give or take $2.50? The cost falls somewhere in the range between $5 and $35.

My single number best guess is _______.

The range in which I would have a good guess goes from _____ to _____.

My guess was _______________ (good / bad)

3. An ink cartridge for an at-home printer costs approximately $15-$40, depending on the size and brand. Can you guess the actual cost to produce a typical cartridge, give or take $0.50? The cost falls somewhere in the range between $0.50 and $6.50.

My single number best guess is _______.

The range in which I would have a good guess goes from _____ to _____.

My guess was _______________ (good / bad)
Let’s take a look at some actual games on The Price Is Right

**Part 2: The Range Game**

Rules: The contestant is shown a prize, then directed to a meter with a range of $600 from the lowest possible guess to the highest possible guess. Within that range is the actual price of the prize. The actual price can occur anywhere with equal probability. A smaller range of $150 (shown in red) will then slowly climb the meter, and the contestant’s task is to stop the rising target range when it contains the actual price of the prize. In doing so correctly, the contestant wins the prize.

In the picture to the right, the lowest possible value is $9,300 and the highest possible value is $9,900. (This setup came from the prize value of a back-yard barbecue island and patio dining set which aired on 10-21-19.)

4. Explain how the probability distribution of the prize value follows a uniform distribution.

5. Using the image above, give the values of a, b, and state the probability density function.

6. Using the image above, what probability does the location of the red range represent?

7. You have no insights as to the true price. No matter where you stop the red range, what is the chance of you winning? Why does it not matter where you stop the range?

8. Do you think knowing the prize ahead of time affects your chances of winning? If so, explain how you believe the probability is affected in terms of the values of a, b, and/or the probability density function. If not, explain why not.

9. Explain how the range game is similar to what we did in Part 1 with the manufacturer’s cost.

10. What is the probability someone would have had a good guess for questions 1, 2, or 3?
Part 3: The Big Wheel

Rules: Three contestants that played the previous mini-games within the show are chosen to play the Showcase Showdown, better known as the “Big Wheel.” The contestants go in order by winnings in the games they just played, from least to greatest. This is because the first player has the lowest probability of winning, and the third player has the highest probability of winning. The wheel consists of the numbers 5 to 100, counting by 5’s, so there are 20 numbers on the wheel, as shown below:

![Big Wheel Image]

Each contestant’s goal is to get closest to $1.00, or 100, without going over. Each contestant is allotted two spins, and must spin if his or her first spin is lower than the final value of a previous contestant. The contestant with the highest score without going over $1.00 in the end is the winner, and moves on to the Final Showcase. In the event of a tie between two or three of the contestants, there is a tiebreaker in which each contestant gets one spin, and the contestant with the highest value for the one spin moves on to the Final Showcase.

When a contestant spins this wheel, it lands on an angle between 0 and 360 degrees with equal probability.

11. Draw an x-y axis on the image above and label the endpoints of each monetary amount with an appropriate angle measure. (Hint: The positive x-axis should cross halfway through the 100 section, and this represents an angle of 0.)

12. Explain how the probability of spinning the big wheel follows a uniform distribution. What are the values of a, b, and the probability density function?
13. What is the probability that when you spin the wheel, it lands somewhere with an angle measure between 125 and 175 degrees? If it lands in this region, what monetary values are possible?

14. What is the probability that your first spin of the wheel lands on the region containing 80? Show how you could determine this with the uniform distribution and the angle of the spin.

15. What is the probability that your first spin lands on 30? On $100? Explain why this probability will always be the same for any value on the wheel.

16. You have a strategy if you’re the first spinning contestant…If you spin a value greater than 0.50, you’ll stop. If you spin $0.50 or less, you take a second spin.
   a. What is the probability that your final total is exactly $1.00?

   b. What is the probability your final total is between $0.90 and $1.00 (inclusive)?

   c. What is the probability you overspin and your total is more than $1.00?

17. Suppose you are the third player to spin and the current total is $0.85. What is the probability you will win or tie the game on your turn?
Activity 9

In-Class Activity: Normal Distributions Weightlifting

Part 1: Read the information below. Then, answer the questions that follow.

Many situations in the real world can be modeled by a Normal Distribution. One example involves weightlifting. Today’s activity uses data compiled over hundreds of thousands of weightlifting attempts, from individuals of varying ability levels who lifted at any official competition.

1) Study the image to the left, which is popular on math-based social media. The amount of wear on the different settings of the weight rack suggests that users’ choice of weight follows a Normal Distribution. Explain how this image supports the idea that weight lifting ability follows a Normal Distribution.

2) For today’s activity, we will be examining two different exercise-related variables. First, we will look at the amount of weight that individuals can deadlift. In the most basic sense, a deadlift refers to the maximum amount that you can lift off of the ground. Second, we will look at the amount of time that you can maintain good form on a plank. Before continuing on with the activity, you will need to determine two data points about yourself. Your instructor will give further instructions on how we are doing this!

   My Deadlift Estimate: _______ lbs

   My Plank Time (or Estimate): ______ secs
Part 2: Study the graphs and information below regarding deadlifts. Then, answer the questions that follow

According to over a hundred thousand data points compiled on OpenPowerLifting.org, Men’s maximum deadlift amount follows an approximately Normal Distribution, with a mean of $\mu = 475 \text{ lb}$ and a standard deviation of $\sigma = 95 \text{ lb}$. Women’s maximum deadlift amount follows an approximately Normal Distribution as well, with a mean of $\mu = 270 \text{ lb}$ and a standard deviation of $\sigma = 80 \text{ lb}$.

3) Take your deadlift estimate from the last page. Using whichever Normal Distribution you prefer, calculate your z-score. Explain what this z-score means, in the context of weightlifting.

4) Using a Standard Normal table, what percent of weightlifters can you outlift?

5) Suppose that you wanted to train until you were in the top 10% of deadlifters. This would mean that you lift at the 90th percentile. Look up what z-score corresponds to the 90th percentile of the Normal Distribution. Then, use the z-score formula to figure out how much weight you would need to lift to be at the 90th percentile.
Part 3: We are going to compile our class plank data. We will assume that plank time also follows a Normal Distribution, and use our class data to approximate the mean and standard deviation of plank time.

Our Class Mean Plank Time: ______ sec

Our Class Standard Deviation of Plank Time: _____ sec

6) Assuming that plank time follows a Normal Distribution, what percentile of plank time does your earlier attempt represent?

7) Assuming that plank time follows a Normal Distribution, what percent of individuals can hold a plank between 40 and 80 seconds?

8) Assuming that plank time follows a Normal Distribution, how long would you need to hold a plank to be at the 20\textsuperscript{th} percentile?

Part 4 – Extra Practice

9) Maximum Squat Weight for female lifters follows an approximately Normal Distribution, with a mean of 230 lbs, and a standard deviation of 40 lbs. What percent of female lifters can squat in between 200 and 300 lbs?

10) Bench Press amounts for male lifters also follow a Normal Distribution, with a mean of 280 pounds. If the top 10\% of men can lift at least 408 pounds, what is the standard deviation of this distribution?
Activity 10

In-Class Activity: Exponential Distributions

March Madness

Part 1: Read the introduction below. Then, read the instructions for the group activity.

In 2021, Baylor upset Gonzaga’s perfect season to win their first ever Men’s Basketball Championship. This year, Baylor is again one of the top teams in the country, but will they be able to pull off back-to-back titles? This hasn’t been accomplished since University of Florida won in 2006 and 2007. Analysts naturally look at many statistics in trying to predict who will win. One statistic of interest is the average time between made baskets. It can be helpful to anticipate how likely a team is to go through a scoring drought, and exponential distributions can help with that! We will first collect our own class data using the activity below.

Group Activity Instructions

● Your group will be competing against the other half of the class to see who can make more shots into your basket over a 4-minute time period.
● Prepare 4-5 “basketballs” by crumpling up individual sheets of paper (no cheating and using multiple sheets at once!)
● Get up and form a separate line for each of the two baskets.
● When you are at the front of the line, you will take a shot and attempt to make a basket. After your shot (whether you make it or not), move to the end of the line, and wait until your next turn.
● Keep taking shots when it’s your turn, until the 4 minutes are up. If you run out of paper balls, you can go grab some off of the floor. The team who made the most baskets wins!

After the Activity...
Record the final score on the blanks below. You will need your group’s number when completing Part 2!

Your Group’s Baskets: ______

Other Group’s Baskets: ______
Part 2: Answer the following questions using your group's results.

Students frequently mix up Exponential settings with Poisson settings. If this was a Poisson setting, we would use the number of baskets as our \( \mu \). In an Exponential Distribution, we don’t care about how many occurrences (baskets) there were. We care about the average **length of time** in between baskets!

1) Is the length of time between baskets a discrete or a continuous random variable? Explain.

2) To determine our \( \mu \) for this problem, we need to calculate the average amount of time in between baskets. This calculation is simple, but it is one that students frequently forget to do! Calculate \( \mu \) using your group’s basket data. Put your answer in **seconds rather than minutes**.

3) Based on the \( \mu \) you calculated above, write out the exponential function that will help us calculate our probabilities.

4) Suppose that your group would have continued shooting after time was called. What is the probability that your group would have made another shot within the next 15 seconds?

5) What is the probability that it would have taken over a minute for your group to make another shot?

6) What is the probability that it would have taken between 30 and 60 seconds?
Part 3: Answer the following questions using the information provided below.

As of 3/4/22, Baylor averaged 28.4 field goals (shots made) per 40-minute game. Their opponents averaged only 23.4 field goals per 40-minute game. Using this information, we can create an exponential equation to model each team’s chances of scoring in the next “x” minutes.

7) Calculate \( \mu \) for Baylor’s offensive exponential distribution. Explain what \( \mu \) tells us, in the context of a basketball game.

8) Calculate \( \mu \) for Baylor’s opponents. Then, include the exponential functions that will allow us to calculate our probabilities in the space below.

Baylor:  

Opponents:

9) What is the probability that Baylor will score within the next 3 minutes? Does your \( \mu \) change when the interval length changes? Why or why not?

10) What is the probability that Baylor’s opponent will not score in the next 2 minutes?
Part 4: Read the description below. Then, select a team from the table.

Major basketball analysts such as Ken Pomeroy (KenPom rankings) have studied the importance of all sorts of basketball statistics on predicting wins. In addition to scoring, a team’s ability to get \textbf{offensive rebounds} (getting the ball back when you miss a shot), and a team’s ability to avoid \textbf{turnovers} (losing the ball without taking a shot) are both extremely important in predicting success.

According to KenPom Rankings as of 3/4/22, the table below shows the top 8 teams in the country. Based on the data, pick the team that you believe is most likely to win the NCAA tournament.

<table>
<thead>
<tr>
<th>Team</th>
<th>Average Field Goals Scored Per Game (High = Good)</th>
<th>Average Field Goals Allowed Per Game (Low = Good)</th>
<th>Average Offensive Rebounds Per Game (High = Good)</th>
<th>Average Turnovers Per Game (Low = Good)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – Gonzaga</td>
<td>33.1</td>
<td>24.3</td>
<td>9.6</td>
<td>11.8</td>
</tr>
<tr>
<td>2 – Baylor</td>
<td>28.4</td>
<td>23.4</td>
<td>12.1</td>
<td>12.8</td>
</tr>
<tr>
<td>3 – Kentucky</td>
<td>30.8</td>
<td>24.4</td>
<td>12.9</td>
<td>11.7</td>
</tr>
<tr>
<td>4 – Arizona</td>
<td>30.7</td>
<td>25.0</td>
<td>11.3</td>
<td>13.3</td>
</tr>
<tr>
<td>5 – Duke</td>
<td>29.7</td>
<td>25.6</td>
<td>10.4</td>
<td>10.3</td>
</tr>
<tr>
<td>6 – Houston</td>
<td>28.5</td>
<td>19.3</td>
<td>13.4</td>
<td>11.1</td>
</tr>
<tr>
<td>7 – Auburn</td>
<td>28.2</td>
<td>22.5</td>
<td>12.0</td>
<td>12.1</td>
</tr>
<tr>
<td>8 – UCLA</td>
<td>28.1</td>
<td>23.1</td>
<td>11.6</td>
<td>9.4</td>
</tr>
</tbody>
</table>

\textbf{Review!} Instead of looking at the time between occurrences (exponential), we can shift to \textit{number} of occurrences in a certain amount of time. This turns the problem into a Poisson Distribution!

11) Take your selected team, and write down the average number of turnovers that they allow in a 40 minute game. This will be our $\mu$ for the Poisson setting.

$$\mu =$$

12) Suppose we look at a 5 minute stretch of your selected team’s game. (Make sure you adjust your $\mu$ accordingly, based on the new time!). \textbf{What is the probability that they will turn the ball over two or more times?} Show your work. You may want to consider using the complement!
APPENDIX D.2:

Control Group Activities

Activity 1

**In-Class Activity: Law of Large Numbers**

**Dice Rolling Game**

Get two dice from your instructor for this activity.

Rules of the game: There are two players, Player A and Player B. Player A will always roll first. Players take turns rolling their die until they have rolled the same number twice (not necessarily in a row). The person who rolls the same number twice loses. Note: Players A and B can roll the same number without the game ending (See Example 3 below.)

Example 1: A rolls 2, B rolls 5, A rolls 2 – Game over, A loses in 3rd round
Example 2: A rolls 2, B rolls 5, A rolls 4, B rolls 3, A rolls 2 – Game over, A loses in 5th round
Example 3: A rolls, 2, B rolls 5, A rolls 5, B rolls 1, A rolls 3, B rolls 5 – Game over, B loses in 6th round

**Part I: Questions to think about with your group...**

How likely is it that the game will end in the third round, as in Example 1?

What is the probability that you roll a 1 on your first roll?

Supposing Player B doesn’t end the game, what is the probability that Player A can keep playing after their second roll of the die? After their third roll?

How does the probability of rolling a “safe” number change each time you roll the dice?

Do you think each player has the same probability of winning the game when player A goes first, or does one player have an advantage over the other?

As a class, let’s vote...Is it better to go first, go second, or does it not matter?
Part II:
With your partner, play the dice rolling game 5 times. Record in the table which player won and the round at which the other player lost. (See example) Player A must take the first turn in all 5 games! (Hint: It may be helpful to have scratch paper to keep track of your rolls.)

<table>
<thead>
<tr>
<th></th>
<th>Player A:</th>
<th>Player B:</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Example given</td>
<td>Loses in round 3</td>
<td>WINNER</td>
</tr>
<tr>
<td>Game 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Game 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What do you notice about the round number in which Player A loses? What about the round number in which Player B loses?

What is the maximum number of rounds possible in any one particular game? Explain your thinking.

Based on your trials, what is the probability of Player A winning? Express your answer as a decimal.

Based on your trials, do you now believe it is better to go first, second, or does it not matter?

Based on your trials, what is the probability of losing in exactly 3 rounds? What is the probability of losing in 7 rounds or more?

Part III: Team up with another group, combine your results and answer these last 3 questions (starting with the words, “Based on your trials”) again...Have your answers changed? Why or why not?
Part IV: Enter the data in your table (from Part II) in the Excel Spreadsheet.

Look at the class results for overall wins:

Wins for Player A: Wins for Player B:

Based on our class data, what is the probability of Player A winning the dice rolling game? Is this a theoretical or experimental probability? How do you know?

What happened to the probability as we used more trials?

Do you think we have enough data to predict the theoretical probability of Player A winning? Why or why not? If not, how many more trials would you recommend?

Now let’s officially answer the question...Is it better to go first, go second, or does it not matter?

Is it more likely for Player A to win 80% of the time if they played only 5 games or if they played 100 games? Explain.
Let’s take a look at the class results for rounds in which the game ends:

What is the most likely number of rounds for a game to end? What is the least likely?

Theoretically, just over 30% of the games played should take 7 or more rounds.

  Scenario A:  5 games were played  
  Scenario B:  50 games were played  
  Scenario C:  500 games were played

In which scenario are you more likely to see that 60% of the games played took 7 or more rounds?

In which scenario are you more likely to see that 30% of the games played took 7 or more rounds?

How does the Law of Large Numbers relate to your answers above?
Activity 2

In-Class Activity: OR Probability
Junk Food

Part 1: Give your best guess on the two following trivia questions related to your professor’s eating/drinking habits

1) What is your professor’s preferred drink with lunch?
   a. Plain Water
   b. Flavored Water
   c. Tea
   d. Coffee
   e. Soda

2) Which of the following fast-food restaurants does your professor eat at most frequently?
   a. Chipotle
   b. Panera
   c. McDonald’s
   d. Subway
   e. Taco Bell

As a class, we are going to vote on what we think the correct answer is. Don’t be influenced by what other people are choosing – stick with your initial answers! Keep track of class votes in the tables below.

<table>
<thead>
<tr>
<th>Country Selected</th>
<th>Plain Water</th>
<th>Flavored Water</th>
<th>Tea</th>
<th>Coffee</th>
<th>Soda</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Votes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attraction Selected</th>
<th>Chipotle</th>
<th>Panera</th>
<th>McDonald’s</th>
<th>Subway</th>
<th>Taco Bell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Votes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Part 2: After we reveal the correct answers to the trivia questions above, use our class data to answer the questions that follow.

3) Let the event C represent getting the correct answer to the beverage question (#1). If we were to select a student at random from our class, what is \( P(C) \)?

4) Let the event A represent getting the correct answer to the fast food question (#2). If we were to select a student at random from our class, what is \( P(A) \)?

5) Suppose that we selected a student at random, and wanted to find \( P(C \cup A) \). Explain in words what this probability would mean.

6) Unfortunately, we cannot just add together \( P(C) \) and \( P(A) \) to get \( P(C \cup A) \). This is because the events C and A are not mutually exclusive! Explain why the events C and A are not mutually exclusive.

7) With “OR” probability, it can be really helpful to draw out the problem using a Venn Diagram. Fill out each blank in the picture below with the correct number of votes from our class.

8) Calculate \( P(C \cup A) \), using either your Venn Diagram above, or the formula shown below.

\[
P(C \cup A) = P(C) + P(A) - P(C \cap A)
\]

9) The complement of an event represents everything not included within the original event. What is the complement of your answer to #8? Explain what this probability represents, in terms of the trivia voting from earlier.
Part 3: Randomly select 10 guilty food pleasures from the list shown below. Then, compare selections with a partner to answer the questions that follow.

You may choose to use a random number table, a random number generator on a calculator, or just randomly list 10 numbers between 01 and 20. The random numbers selected are:

_____, _____, _____, _____, _____, _____, _____, _____, _____, _____

<table>
<thead>
<tr>
<th>01 - Ice Cream</th>
<th>02 - Potato Chips</th>
<th>03 - Cookies</th>
<th>04 - French Fries</th>
<th>05 - Macaroni &amp; Cheese</th>
</tr>
</thead>
<tbody>
<tr>
<td>06 - Chocolate / Chocolate Bar</td>
<td>07 - Donuts</td>
<td>08 - Brownies</td>
<td>09 - Cheesecake</td>
<td>10 - Pretzels</td>
</tr>
<tr>
<td>11 - Cupcakes</td>
<td>12 - Pizza / Pizza Rolls</td>
<td>13 - Milk Shakes</td>
<td>14 - Movie Theater Popcorn</td>
<td>15 - Hot Dogs</td>
</tr>
<tr>
<td>16 - Hamburger / Cheeseburger</td>
<td>17 - Nachos</td>
<td>18 - Soda</td>
<td>19 - Candy</td>
<td>20 - Chicken Nuggets</td>
</tr>
</tbody>
</table>

Let P(Y) represent the probability that a food above is on your list, and let P(Z) represent the probability that a food is on your partner’s list. Obviously, these probabilities are both equal to 10/20, since you were instructed to select 10 sites.

10) Compare and discuss your selections with a partner. Place check marks on the table above to represent foods that you both selected. Then, find the elements in set \( Y \cap Z \) and \( P(Y \cap Z) \).

11) Using your answer to #10, calculate \( P(Y \cup Z) \).

12) Find \( P(Y^c \cap Z^c) \). Explain what this probability represents in the context of the guilty pleasure foods.
Part 4: Use the following table to get additional practice with “OR” probability.

An internet search found a list of the top 50 indulgent foods. Let the event $S =$ the food is sweet, and the event $F =$ the food was fried. Results are compiled in the table below.

<table>
<thead>
<tr>
<th></th>
<th>Sweet</th>
<th>Not Sweet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fried</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Not Fried</td>
<td>26</td>
<td>7</td>
</tr>
</tbody>
</table>

13) Find $P(S \cap F)$

14) Find $P(S \cup F)$

15) Find $P(S \cup F^C)$
Activity 3

In-Class Activity: Tree Diagrams
Widgets & Quality Control

Part 1: Carefully read the information below to familiarize yourself with today’s lesson context.

A company is selling “widgets” and must examine them for any defects before releasing their product to the general public. The company has two scanning machines (A and B) that are available to inspect widgets. The scanning machines inspect each widget individually and attempt to catch and remove any defective widgets that are present.

- Machine A will catch a defective widget 91% of the time (p = 0.91).
- Machine B will catch a defective widget 82% of the time (p = 0.82).

Part 2: Use the probability rates given above to answer the following questions.

1) Suppose that the company decides to use only Machine A for inspection. If a defective widget is scanned by Machine A, what is the probability that it will not be caught?

2) Below is a tree diagram showing two defective widgets passing through Machine A for inspection. What is the probability that at least 1 of the defective widgets will not be caught? Show two ways to arrive at your answer.
3) Expand the tree diagram to include 3 defective widgets which pass through Machine A during the day. What is the probability that at least 1 of the defective widgets will not be caught?

4) Suppose that the company decides to use only Machine A for inspection. Assume that over the course of a day, 20 defective widgets pass through inspection. What is the probability that at least 1 of the defective widgets will not be caught?

5) Suppose that the company decides to use only Machine B for inspection. If a defective widget is scanned by Machine B, what is the probability that it will not be caught?

6) Suppose that the company decides to use only Machine B for inspection. Assume that over the course of a day, 20 defective widgets pass through inspection. What is the probability that at least 1 of the defective widgets will not be caught?
Part 3: By scanning each widget with both machines, the company can drastically reduce the chances of a defective widget evading detection. Answer the following questions, using a tree diagram to assist in calculations.

7) The tree diagram below represents the possible outcomes when every widget is inspected by both machines. Fill out the diagram, using the rates provided on the front page. Then, calculate the probability for each of the 4 outcomes.

8) Based on your tree diagram above, what is the probability that a defective widget will be missed by both Machine A and Machine B? Show your work.

9) Based on the probability you calculated in #8, how many times lower is the risk of missing a defective widget, when compared to using Machine A alone?

Part 4: For additional practice, calculate the probability below by creating a tree diagram.

10) For lunch, you always eat either a frozen burrito or a cup of ramen. While shopping, you buy 4 burritos, and 7 ramen cups. For three days in a row, you randomly choose either a burrito or ramen for lunch. What is the probability that you will eat at most 1 burrito during the 3 lunches?
Activity 4

In-Class Activity: Conditional Probability
Earthworms

Part 1: Read this excerpt from “Tiny Earthworm’s Big Impact”.
https://www.sciencenewsforstudents.org/article/tiny-earthworms-big-impact Then answer the questions about the article.

Gardeners tend to like earthworms because they mix the soil, loosening it and moving nutrients around. Earthworms even shred leftover plant parts into smaller fragments eaten by microorganisms. In these ways, earthworms can improve and enrich the soil, allowing garden and certain crop plants to grow better. But many American scientists are coming to view some earthworms as enemies.

In the 1600s, European settlers brought European earthworms to North America. At that time, the continent’s northern forests had no soil-mixing earthworms. If any had once existed there, they were likely very different from the European species. And they would have been wiped out during the glacial period that ended 11,000 years ago.

Today, in these forests, legions of earthworms blend soil with plant scraps like fallen leaves and twigs. And that mixing has proven disastrous for the complex network of soil, water, plants and animals — the ecosystem — that developed over thousands of years without earthworms. Since arriving in North America, invasive earthworms have changed the landscape, assisted other non-native species with getting a foothold, and competed with native species.

A study was done and a variety of scientists, gardeners, and outdoor enthusiasts were asked about their opinion on earthworms. Half of the 240 people interviewed lived in a region inhabited by earthworms and the other half lived in an earthworm-free habitat. Does living in the presence of earthworms affect your opinion of them? It appears so! Only 10% of respondents thought earthworms did more damage than good when they lived in a region inhabited by earthworms, but this increased to nearly 30% when the respondent lived in an earthworm-free region.
1) For the rest of this section, let’s use $E$ to denote the event of living in a region with earthworms, $F$ to denote a favorable opinion of earthworms, and $U$ to denote an unfavorable opinion. What statistics or probabilities did you see on the previous page and what symbolic notation or formula could you use with them?

2) Fill in the probability table below. You may need to do a few calculations to complete it. *(Hint: Do the 10% and 30% have a place on the table, or are they describing something else?)*

<table>
<thead>
<tr>
<th></th>
<th>$E$</th>
<th>$E^C$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3) Use the information in the table to answer the following questions:

A.) If one of these 240 people were randomly chosen, what is the probability that they had a favorable opinion of earthworms and did not live in an earthworm-inhabited region?

B.) Is this answer to Question A the same as $P(F) \times P(E^C)$? Why or why not?

C.) Are the events of having a favorable opinion and not living around earthworms independent, dependent, and/or mutually exclusive? Explain your choice(s).

D.) Out of the 240 people in the survey, how many of them actually had an unfavorable opinion of earthworms?

E.) Given that a person lives in an earthworm inhabited region, what is the probability that they have a favorable opinion of earthworms?
Part 2: 1,210 participants took an online survey to determine how much they knew about worms. 60% of the participants were women (W) and 40% of the participants were men (M). Other results from the study were:

- 15% of women and 20% of men said they knew worms become paralyzed if exposed to light for too long (A).
- 50% of women and 35% of men said that they knew worms could regenerate segments of their bodies (R).
- 60% of women and 72% of men said they knew that worms breathed through their skin and had no lungs (L).
- 10% of women and 12% of men said they knew worms were cold-blooded and had five hearts (C).
- 75% of women and 80% of men said they knew worms were hermaphrodites (both male and female). (H)

Select one of these bullet points and create a probability table based on the data.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Enter your letter indicated in your chosen bullet point in the questions below, determine the probabilities, state if they are joint, marginal, or neither, and then give a verbal description of what the probability indicates:

\[ P(\text{__}) \]

\[ P(W \cap \text{__}) \]

\[ P(M \cup \text{__}^c) \]

\[ P(W|\text{__}) \]
A mining company collects and scans large rocks, in an effort to obtain Mineral “Z” which is found in a small percentage of rocks in the quarry. The company has a machine that can scan the rocks for Mineral Z, but it takes time and effort to load each rock into the scanner. Suppose that a cart containing 12 large rocks arrives for scanning. Assume that 10% of large rocks in the mine contain this valuable Mineral Z, and presence of “Z” is independent from rock to rock. The company is considering two different strategies for scanning these 12 rocks:

- **Method A:** Load each rock into the scanner individually, and scan each of the 12 rocks separately.
- **Method B:** Run the entire cart of 12 rocks through the scanner first.
  - If the full-cart scan detects any traces of Mineral “Z”, the entire cart is brought back out. Then, each of the 12 rocks will need to be loaded into the machine and re-scanned individually.
  - If the full-cart scan does **not** detect any traces of Mineral “Z”, the whole batch of rocks can be dumped without any more lifting or scanning.

a. Let the random variable $X$ represent the total number of scans that a cart of 12 rocks will require using Method B (batch scanning method). Determine the probability distribution of $X$.

<table>
<thead>
<tr>
<th>Number of Scans (X)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. Is the variable $X$ discrete or continuous? Explain.

c. If the company implements Method B on one cart of 12 rocks, what is the probability that the number of scans needed will be smaller than it would be with Method A?
d. Determine the expected value of X.

e. Interpret this expected value in the context of the problem.

f. If you had to scan thousands of carts of 12 rocks, which method - A or B - is better? Why?

Now consider Method C: We will first randomly divide the 12 rocks into two carts with 6 rocks each. Each of the carts will be wheeled through the scanning machine separately, using a similar process to above.

- As before, if a scan of a cart detects Mineral “Z”, the entire cart will need to be brought back, and all 6 rocks within the cart must be loaded into the scanner individually.
- As before, if a scan of a cart does not detect Mineral “Z”, the entire cart of rocks can be discarded without any further scanning.

g. Let the random variable Y represent the total number of tests needed for a batch of 12 rocks with Method C (12 rocks separated into batches of 6). Determine the probability distribution of Y.

<table>
<thead>
<tr>
<th>Number of Tests (Y)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

h. If you implement Method C once, what is the probability that it will require fewer scans than Method A?

i. Determine the expected value of Y. Is Method C better than Method B? Why?
j. Calculate and interpret the standard deviation of $Y$.

k. Discuss the following questions with your group:

This CEO of the mining company has decided to only batch test rocks with no visual evidence of mineral Z. Why do you think the CEO made this rule?

If the mining company begins to operate in a new mine and finds that Mineral Z is much more prevalent in those rocks, batch testing begins to make less sense. Explain why this is the case. (Hint...What would happen if Mineral Z appears in 50% of the rocks in the new mine?)

How does the number of rocks included in the batch have an impact on the results from your batch testing?
Activity 6

In-Class Activity: The Binomial Distribution
Rock Paper Scissors

Rock Paper Scissors is a hand game in which players make one of three gestures with their hands (as shown in the image above). The arrows in the diagram indicate which object “wins” over the other. For instance, in a battle between Paper and Rock, Paper would win. But in a battle between Paper and Scissors, Scissors would win.

**Part 1:** Find a partner, not in your typical working group. Play 5 or more games of Rock Paper Scissors where you record the results on your page and 5 or more games of Rock Paper Scissors where the other person records the results on their page. Tally your wins, losses, and ties in the table below.

<table>
<thead>
<tr>
<th>Wins</th>
<th>Losses</th>
<th>Ties</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I won on _____ out of _____ games.

**Part 2:** Join with your group and fill in the table below:

<table>
<thead>
<tr>
<th>Group Member</th>
<th># of Wins</th>
<th>Total # of Games</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What was the theoretical probability of winning one game?

What was the experimental probability of winning one game?
**Part 3:** Let $X$ be a random variable that represents the number of wins out of all the games for members in your group. What type of distribution does $X$ have? Justify your response, indicating values for any of the variables that describe the distribution. (If needed, use the value from the theoretical probability rather than the experimental probability.)

**Part 4:** Using the distribution identified above, answer the following questions:

1. What is the probability that your group would have won all their games?

2. What is the probability that another group, playing the same amount as your group did, would have more wins than your group? (Hint: You may want to use technology to help you here.)

3. For the number of games your group played, what is the expected number of wins, give or take how much?

4. Is winning this expected number actually a realistic possibility? Why or why not?
Activity 7

In-Class Activity: Poisson Distribution

Leaking Pipes

Part 1: Read the introduction below. Then, read the instructions for the group activity.

Poisson Distributions can be used to predict the number of failures that will occur in a certain interval of time or a length of something. This can include events such as power outages, accidents on a highway, or leaks along a pipeline. In this lesson, we will look at leaks along a supply pipe. Since we cannot simulate this in person, we will use a word game to represent failures, or leaks, in a set amount of time.

Group Activity Instructions

● Suppose all members of your group represent a particular portion of our pipeline.
● You must not speak words that start with the letter “T”. Every time you do, it represents a “leak” that occurs along the pipeline. Each minute you speak will represent a mile of pipeline.
● For 4 minutes, group members will go around in a circle, describing something good that they recently had to eat. Go into excessive detail – when was it? Where were you? What made it so good? While each person is talking, the others should closely listen and keep track of how many times the speaker used words that start with “T”. You should aim to talk for a full minute each.

After the Activity...

● Once everyone has had a turn, let each person in your group know how many times they may have had a “leak” in their 1 minute of talking (that is, in 1 mile of pipeline).
● Take the average number of leaks for your group. This value will serve as your group’s expected number of leaks per mile, or $\mu$.

\[
\text{My Group's } \mu: _______
\]

Suppose that the probability of a leak was the same over any equal interval of pipeline. This situation can be modeled by a Poisson Distribution! Use your group’s $\mu$ to calculate the probabilities that follow.
Part 2: Answer the following questions using your group’s $\mu$

1) In a 1-mile stretch of pipeline, what is the probability that exactly 6 leaks will occur?

2) In a 1-mile stretch of pipeline, what is the probability that more than 2 leaks will occur?

3) In a 4-mile stretch of pipeline, what is the probability that exactly 30 leaks will occur? Why is this number so small?

Part 3: Answer the following questions using the information provided below.

Rather than looking at a portion of pipeline, let’s take a look at the pipeline as a whole, and focus on how often the leaks are occurring. Suppose that you are in charge of quickly finding and fixing leaks on the Dakota Access Pipeline. During its first six months in operation in 2017, the Dakota Access Pipeline reported 5 leaks. If we assume this failure rate continued, then for any six-month period of time, $\mu = 5$. This would mean that we can expect an average of 5 leaks over a 6-month period.

4) In a 6-month period, what is the probability that there will be no leaks? Use the Poisson Formula.

5) In a 6-month period, what is the probability that there will be exactly 3 leaks? Use the Poisson Formula.

6) In a 6-month period, what is the probability that there will be either 1 or 2 leaks? Use the Poisson Table Provided in Question 9.
7) Over a 3-month period, what is the probability that there will be exactly 4 leaks? Use the Poisson Formula.

8) Over a 1-year (not month!!) period, what is the probability that there will not be any leaks? Use the Poisson Formula.

9) Over a 1-year period, what is the probability that there will be at least 2 leaks? Use the Poisson Table provided below.

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>1.0</th>
<th>2.0</th>
<th>2.5</th>
<th>3.0</th>
<th>4.0</th>
<th>5.0</th>
<th>6.0</th>
<th>10.0</th>
<th>12.0</th>
<th>20.0</th>
<th>24.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.67</td>
<td>0.15</td>
<td>0.08</td>
<td>0.04</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1</td>
<td>0.36</td>
<td>0.27</td>
<td>0.20</td>
<td>0.14</td>
<td>0.10</td>
<td>0.07</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>0.18</td>
<td>0.26</td>
<td>0.20</td>
<td>0.16</td>
<td>0.11</td>
<td>0.08</td>
<td>0.05</td>
<td>0.03</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>0.09</td>
<td>0.21</td>
<td>0.16</td>
<td>0.12</td>
<td>0.09</td>
<td>0.07</td>
<td>0.05</td>
<td>0.03</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>0.05</td>
<td>0.14</td>
<td>0.12</td>
<td>0.09</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>0.03</td>
<td>0.10</td>
<td>0.09</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>0.02</td>
<td>0.08</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>0.01</td>
<td>0.07</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>8</td>
<td>0.00</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>9</td>
<td>0.00</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>11</td>
<td>0.00</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>12</td>
<td>0.00</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>13</td>
<td>0.00</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>14</td>
<td>0.00</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>15</td>
<td>0.00</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.00</td>
</tr>
</tbody>
</table>

10) Your boss promises that if you can go 2 years (not months!) without a leak occurring, you will receive a huge bonus. How likely is this to occur? Justify using information related to the Poisson Distribution.
Part 4 (optional, if time): In reality, the operators of the DAPL slowed the leak rate down after a lot of bad press. After 3 years in operation, 10 total leaks had been reported.

11) Come to a consensus within your group of what would be an appropriate value of $\mu$ to use to measure the failure rate of the DAPL pipeline. Should you take those disastrous first 6 months into account or not?

12) Using your value of $\mu$ (above), determine the probability that DAPL will have another leak before this calendar year is over? (Hint: How many more months are left in the year? What value of $\mu$ would be needed for this problem?)

13) Discuss the practical implications of this calculation with your group. Should the Standing Rock Sioux tribe be worried? Does the length of pipeline make a difference? (It is 1,172 miles from the Dakotas down to Illinois). Would a different calculation, such as leaks per mile, help you to decide if the DAPL pipeline leaks are worrisome?
Activity 8

In-Class Activity: The Uniform Distribution
Number Games

Part 1: Guess that value!

1. The variable A has some value in the range between $5 and $125. Can you guess the
correct value, give or take 10?

| 5  | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 70 | 75 | 80 | 85 | 90 | 95 | 100 | 105 | 110 | 115 | 120 | 125 |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|

   My single number best guess is _______.
   The range in which I would have a good guess goes from ______ to _______.
   My guess was _______________ (good / bad)

2. The variable B has some value in the range between 5 and 35. Can you guess the
correct value, give or take 2.5?

<table>
<thead>
<tr>
<th>5:00</th>
<th>7:00</th>
<th>9:00</th>
<th>11:00</th>
<th>13:00</th>
<th>15:00</th>
<th>17:00</th>
<th>19:00</th>
<th>21:00</th>
<th>23:00</th>
<th>25:00</th>
<th>27:00</th>
<th>29:00</th>
<th>31:00</th>
<th>33:00</th>
<th>35:00</th>
</tr>
</thead>
</table>

   My single number best guess is _______.
   The range in which I would have a good guess goes from ______ to _______.
   My guess was _______________ (good / bad)

3. The variable C has some value in the range between 0.5 and 6.5. Can you guess the
correct value, give or take 0.50?

<table>
<thead>
<tr>
<th>0.50</th>
<th>1.00</th>
<th>1.50</th>
<th>2.00</th>
<th>2.50</th>
<th>3.00</th>
<th>3.50</th>
<th>4.00</th>
<th>4.50</th>
<th>5.00</th>
<th>5.50</th>
<th>6.00</th>
<th>6.50</th>
</tr>
</thead>
</table>

   My single number best guess is _______.
   The range in which I would have a good guess goes from ______ to _______.
   My guess was _______________ (good / bad)
Part 2: The Sliding Range Game

Rules: There is a range of values in which a winning number falls. This winning number is equally likely to fall anywhere in the range. You are given a red window, which is 150 wide, and instructed to slide it somewhere on the range where you believe the winning number will be located. If the number falls in the red window, you win the game.

4. Explain how the probability distribution of the prize value follows a uniform distribution.

5. Using the image above, give the values of a, b, and state the probability density function.

6. Using the image above, what probability does the location of the red window represent?

7. You have no insights as to the winning value. No matter where you slide the red window, what is the chance of you winning? Why does it not matter where you slide the range?

8. Do you think having some insight about the winning value affects your chances of winning? If so, explain how you believe the probability is affected in terms of the values of a, b, and/or the probability density function. If not, explain why not.

9. Explain how the sliding range same is similar to what we did in Part 1.

10. What is the probability someone would have had a good guess for questions 1, 2, or 3?
Part 3: Spin The Wheel

Rules: The wheel consists of the numbers 1 to 20, as shown to the right:

Three players play this game by spinning the wheel. Each player's goal is to get closest to 20, without going over. Each player is allotted two spins, and must spin a second time if his or her first spin is lower than the final value of a previous player. The values of the two spins will then be added together for a final score. The player with the highest final score without going over 20 in the end is the winner.

In the event of a tie between two or three of the players, there is a tiebreaker in which each player gets one spin, and the player with the highest value for the one spin is the official winner.

When a player spins this wheel, it lands on an angle between 0 and 360 degrees with equal probability.

11. Draw an x-y axis on the image above and label the endpoints of sector with an appropriate angle measure. (Hint: The positive x-axis should cross between the 2 and the 3, and this represents an angle of 0.)

12. Explain how the probability of spinning the wheel follows a uniform distribution. What are the values of a, b, and the probability density function?

13. What is the probability that when you spin the wheel, it lands somewhere with an angle measure between 125 and 175 degrees? If it lands in this region, what numerical values are possible?

14. What is the probability that your first spin of the wheel lands on the region containing 16? Show how you could determine this with the uniform distribution and the angle of the spin.
15. What is the probability that your first spin lands on 6? On 20? Explain why this probability will always be the same for any value on the wheel.

16. You have a strategy if you’re the first player... If you spin a value greater than 10, you’ll stop. If you spin 10 or less, you take a second spin.
   a. What is the probability that your final total is exactly 20?
   
   b. What is the probability your final total is between 18 and 20 (inclusive)?

   c. What is the probability you overspin and your total is more than 20?

17. Suppose you are the third player to spin and the current total is 17. What is the probability you will win or tie the game on your turn?
Activity 9

In-Class Activity: Normal Distribution

Part 1: Read the information below. Then, answer the questions that follow.

Many situations in the real world can be modeled by a Normal Distribution. IQ Scores, distributions of height, and machine dispensing amounts are just a few examples. In our lesson today, we are going to look at the distribution of weight for newborn infants.

1) Why does it make sense that the weight distribution of babies would be Normally distributed? Explain what this would mean in terms of the frequency of different newborn weights.

2) How much do you think that the average newborn weighs? Express your answer in pounds, using a decimal if needed. Don’t look up the answer, or look on to the next page of the activity!

   My Estimate: _______ lbs

3) Compare answers with at least two other people in class. Once you have done so, you may adjust your estimate if you’d like (but you don’t have to!). You’ll be using this estimate on the next page.

   My Updated Estimate: _______ lbs

We also will be collecting class data for later in the activity. On your instructor’s desk, you will see a number of dice. Grab 2 dice, roll them, and find the sum of your two dice. Your job is to continue rolling pairs of dice until you roll a sum of 7. Have a partner time how long (in seconds, NOT in terms of the number of rolls) it takes until you roll your 7, and record this value below. We will compile this class data and use it in Part 3 of our activity.

   My Time to Roll a Sum of 7: _______ secs
Part 2: Use the information provided below about newborn birthweight to answer the questions that follow.

Although slight differences exist between sexes, birthweight of newborns follow an approximately Normal Distribution, with a mean of $\mu = 7.73$ lb and a standard deviation of $\sigma = 0.85$ lb.

4) Take your birthweight estimate from last page. Calculate the z-score for a newborn that weighs as much as your estimate. Explain what this z-score means, in the context of baby weight.

5) Using a Standard Normal table, what percent of babies weigh less than your estimate?

6) Suppose that you wanted to determine the cutoff for the heaviest 10% of newborns. This would require finding the 90th percentile for birth weight. Look up what z-score corresponds to the 90th percentile of the Normal Distribution. Then, use the z-score formula to figure out how much a baby at the 90th percentile would weigh.

Part 3: We are going to compile our class dice data. We will assume that the time it takes to roll a sum of 7 also follows a Normal Distribution, and use our class data to approximate the mean and standard deviation of this variable.

Our Class Mean Time to Roll a 7: ______ sec

Our Class Standard Deviation of Time to Roll a 7: _____ sec

7) Assuming that 7 rolling time follows a Normal Distribution, what percentile of roll time does your own attempt represent?
8) Assuming that 7 rolling time follows a Normal Distribution, what percent of individuals will take between 15 and 30 seconds to roll a 7?

9) Again, assume that 7 rolling time follows a normal distribution. If you roll a 7 on your first attempt (in other words, you were atypically fast), would you be at a lower percentile or a higher percentile? Explain your thinking.

10) Assuming that 7 rolling time follows a Normal Distribution, how quickly would you need to roll a 7 to be at the 20th percentile?

Part 4 – Extra Practice

11) The distribution of human pregnancy duration follows an approximately Normal Distribution with a mean of 266 days, and a standard deviation of 16 days. What proportion of all pregnancies will last between 240 and 270 days?

12) Newborn length follows an approximately Normal Distribution with a mean of 49.98 cm. If the top 25% of newborns are at least 51.77 cm long, what is the standard deviation for newborn length?
Activity 10

In-Class Activity: Exponential Distributions

WWII Planes

Part 1: Read the introduction below. Then, read the instructions for the group activity.

Landing on an aircraft carrier is perhaps one of the toughest feats in all of aviation. In fact, studies have shown that pilots are more anxious about a night-time carrier landing than they are about combat. Today, there are a number of systems in place to help a pilot get down safely, but during World War II, it was a lot harder.

Just like today, there was a landing signals officer (LSO) responsible for the safe recovery of carrier aircraft, but they didn’t have the modern tools available now. No, this guy had to use paddles and hand gestures to get a planes, like the F6F Hellcat or SBD Dauntless, back on the boat safely. The carriers back then didn’t have angled decks, either. They didn’t have modern radios like the ones we enjoy today. The primitive radios were only able to send Morse code, and sending code isn’t very conducive to getting urgent messages to pilots quickly and clearly. Instead, the LSO stood in a very exposed position and used a pair of paddles to send the pilot signals and guide them into a safe landing.

In order to land safely, pilots needed to be flying in line with the carrier, at an appropriate angle of descent and at an appropriate speed. Because the planes had a short distance in which to land and come to a stop, aircraft were fitted with retractable hooks that engaged transverse wires on the deck, braking them to a quick stop. Pilots who landed too long would miss the wire and fly off the end of the ship.

Group Activity Instructions

- Your group will be competing against the other half of the class to see who can land more planes on a “carrier” over a 4-minute time period.
- Prepare 4-5 paper airplanes by using individual sheets of paper. You may style them however you like, but it needs to be done expediently.
- Get up and form a separate line for each of the two carriers.
- When you are at the front of the line, you will fly your plane and attempt to make a landing. After your attempt (whether you make it or not), move to the end of the line, and wait until your next turn.
- Keep making landing attempts when it’s your turn, until the 4 minutes are up. If you run out of paper planes, you can go grab some off of the floor. The team who made the most landings wins!
After the Activity...
Record the final score on the blanks below. You will need your group’s number when completing Part 2!

Your Group’s Landings: ______

Other Group’s Landings: ______

Part 2: Answer the following questions using your group’s results.
Students frequently mix up Exponential settings with Poisson settings. If this was a Poisson setting, we would use the number of landings as our \( \mu \). In an Exponential Distribution, we don’t care about how many occurrences (landings) there were. We care about the average length of time in between landings!

1) Is the length of time between landings a discrete or a continuous random variable? Explain.

2) To determine our \( \mu \) for this problem, we need to calculate the average amount of time in between landings. This calculation is simple, but it is one that students frequently forget to do! Calculate \( \mu \) using your group’s landing data. Put your answer in seconds rather than minutes.

3) Based on the \( \mu \) you calculated above, write out the exponential function that will help us calculate our probabilities.

4) Suppose that your group would have continued flying planes after time was called. What is the probability that your group would have made another landing within the next 15 seconds?
5) What is the probability that it would have taken over a minute for your group to make another landing?

6) What is the probability that it would have taken between 30 and 60 seconds?

Part 3: Answer the following questions using the information provided below.
During the last 20 months of WWII, each plane flew an average of 15.4 times per month. (Source: [https://www.history.navy.mil/content/dam/nhhc/research/histories/ naval-aviation/aviation-monographs/nasc.pdf](https://www.history.navy.mil/content/dam/nhhc/research/histories/ naval-aviation/aviation-monographs/nasc.pdf)) Let’s assume we are living within these last 20 months of the war. Using this information, we can create an exponential equation to model each plane’s chances of flying in the next “x” days. (Assume there are 30 days in each month.)

7) Calculate \( \mu \) for a WWII plane’s flight exponential distribution. Explain what \( \mu \) tells us, in the context of the plane and/or flight. Then write the exponential function that will allow us to calculate probabilities.

8) What is the probability that a plane will have a flight within the next 3 days? Does your \( \mu \) change when the interval length changes? Why or why not?

9) What is the probability a plane will be grounded for more than a week?
While in flight, accurately dropping a bomb on target was also a tricky procedure during WWII, due to the lack of on-board computers and advanced technology that we enjoy today. During the last 20 months of WWII, our Pacific fleet dropped an average of 1687.5 tons of bombs on target per month. (Source: https://www.history.navy.mil/content/dam/nhhc/research/histories/naval-aviation/aviation-monographs/nasc.pdf) Let’s continue to assume we are living within these last 20 months of the war, and let’s now also assume that each bomb weighed 1 ton. (In reality, most WWII bombs weighed between 100 and 3,000 pounds, while others like the British Grand Slam were as large as 22,000 pounds!) Using this information, we can create an exponential equation to model the chances of a bomb landing on its target in the next “x” hours. (Assume there are 30 days in each month.)

10) Calculate $\mu$ for a Pacific bomb’s exponential distribution. Explain what $\mu$ tells us, in the context of a bomb. Then write the exponential function that will allow us to calculate probabilities.

11) What is the probability that a bomb lands on target in the Pacific within the next hour?

12) What is the probability that 30 minutes pass without a bomb landing on target?

Part 4:
The table below shows 8 of major war actions during the last 20 months of WWII. Based on the data, pick the one that interests you the most.

<table>
<thead>
<tr>
<th>Major War Actions</th>
<th>Date (days)</th>
<th>Flights</th>
<th>Tons of Bombs on target (assume 1 ton = 1 bomb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marshalls Campaign</td>
<td>1/30/44 to 2/23/44 (25 days)</td>
<td>7,387</td>
<td>2,261</td>
</tr>
<tr>
<td>Marianas Campaign</td>
<td>6/11/44 to 8/8/44 (59 days)</td>
<td>22,432</td>
<td>7,090</td>
</tr>
<tr>
<td>Leyte Campaign</td>
<td>10/10/44 to 11/25/44 (47 days)</td>
<td>15,327</td>
<td>4,853</td>
</tr>
<tr>
<td>Mindoro Campaign</td>
<td>12/13/44 to 12/17/44 (5 days)</td>
<td>2,062</td>
<td>333</td>
</tr>
<tr>
<td>Lingayen Campaign</td>
<td>1/3/45 to 1/30/45 (28 days)</td>
<td>8,637</td>
<td>2,308</td>
</tr>
<tr>
<td>Iwo Jima Campaign</td>
<td>2/16/45 to 3/8/45 (21 days)</td>
<td>8,091</td>
<td>1,691</td>
</tr>
<tr>
<td>Okinawa Campaign</td>
<td>3/18/45 to 6/22/45 (97 days)</td>
<td>40,157</td>
<td>12,888</td>
</tr>
<tr>
<td>Assault on Japan</td>
<td>7/10/45 to 8/15/45 (37 days)</td>
<td>12,153</td>
<td>4,382</td>
</tr>
</tbody>
</table>
Review! Instead of looking at the time between occurrences (exponential), we can shift to *number* of occurrences in a certain amount of time. This turns the problem into a Poisson Distribution!

13) Take your selected military action, and write down the average number of bombs on target per day. This will be our $\mu$ for the Poisson setting.

$$\mu =$$

14) Suppose we look at a 2-hour stretch of your selected military action. (Make sure you adjust your $\mu$ accordingly, based on the new time!). **What is the probability that two or more bombs will land on target?** Show your work. You may want to consider using the complement!
**APPENDIX E.1**

Demographic Tables

**Table E.1.1**

*Demographic Data on Educational Level for Treatment and Control Groups*

<table>
<thead>
<tr>
<th></th>
<th>Freshman</th>
<th>Sophomore</th>
<th>Junior</th>
<th>Senior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Group</td>
<td>19% ((n=8))</td>
<td>23% ((n=10))</td>
<td>40% ((n=17))</td>
<td>19% ((n=8))</td>
</tr>
<tr>
<td>Control Group</td>
<td>14% ((n=4))</td>
<td>21% ((n=6))</td>
<td>45% ((n=13))</td>
<td>21% ((n=6))</td>
</tr>
</tbody>
</table>

**Table E.1.2**

*Demographic Data on Gender for Treatment and Control Groups*

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>Nonbinary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Group</td>
<td>44% ((n=19))</td>
<td>53% ((n=23))</td>
<td>2% ((n=1))</td>
</tr>
<tr>
<td>Control Group</td>
<td>45% ((n=13))</td>
<td>55% ((n=16))</td>
<td>0% ((n=0))</td>
</tr>
</tbody>
</table>

**Table E.1.3**

*Demographic Data on Citizenship for Treatment and Control Groups*

<table>
<thead>
<tr>
<th></th>
<th>US Citizen</th>
<th>Foreign Student</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Group</td>
<td>91% ((n=39))</td>
<td>5% ((n=2))</td>
<td>5% ((n=2))</td>
</tr>
<tr>
<td>Control Group</td>
<td>100% ((n=29))</td>
<td>0% ((n=0))</td>
<td>0% ((n=0))</td>
</tr>
</tbody>
</table>
Table E.1.4

Demographic Data on Choice of Major for Treatment and Control Groups

<table>
<thead>
<tr>
<th></th>
<th>Biology</th>
<th>Business</th>
<th>Economics</th>
<th>Education</th>
<th>Engineering</th>
<th>Nursing</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>2%</td>
<td>63%</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
<td>21%</td>
<td>7%</td>
</tr>
<tr>
<td>(n=1)</td>
<td>(n=27)</td>
<td>(n=1)</td>
<td>(n=1)</td>
<td>(n=1)</td>
<td>(n=9)</td>
<td>(n=3)</td>
<td></td>
</tr>
<tr>
<td><strong>Control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>0%</td>
<td>76%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>14%</td>
<td>10%</td>
</tr>
<tr>
<td>(n=0)</td>
<td>(n=22)</td>
<td>(n=0)</td>
<td>(n=0)</td>
<td>(n=0)</td>
<td>(n=4)</td>
<td>(n=3)</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX E.2

Attendance Records

Table E.2

<table>
<thead>
<tr>
<th>Attendance Records for Treatment and Control Groups</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Number of Intervention Activities Attended</strong></td>
<td><strong>Number of Treatment Group Students</strong></td>
<td><strong>Number of Control Group Students</strong></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>
## APPENDIX E.3

Final Code Book

<table>
<thead>
<tr>
<th>Factors Affecting Comprehension</th>
<th>Description</th>
<th>In Vivo Coding</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Elements of Instructional Process</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Amount of Material</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A lot of Material</td>
<td>The amount of material in the unit.</td>
<td>“It’s too much.”</td>
<td>18</td>
</tr>
<tr>
<td>Difficult</td>
<td>The level of complexity.</td>
<td>“A lot more being on.”</td>
<td>6</td>
</tr>
<tr>
<td>Formulas</td>
<td>The formulas in the unit affected their comprehension.</td>
<td>“Symbols being reused in different ways.”</td>
<td>4</td>
</tr>
<tr>
<td>Formatting of Questions</td>
<td>The way that the questions are detailed.</td>
<td>“Made the questions bold so it stood out.”</td>
<td>1</td>
</tr>
<tr>
<td>Probability Content &amp; Topics</td>
<td>The different probability unit topics.</td>
<td>“Distinguish between which ones I would be using.”</td>
<td>7</td>
</tr>
<tr>
<td>Justification</td>
<td>Providing the reasoning.</td>
<td>“This is why we do it.”</td>
<td>1</td>
</tr>
<tr>
<td><strong>Class Structure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flipped Classroom</td>
<td>The lecture materials are outside of the class time.</td>
<td>“The Flipped classroom setting was not useful.”</td>
<td>11</td>
</tr>
<tr>
<td>Groups</td>
<td>The group work during the in-class activity</td>
<td>“Groups activities are discouraging.”</td>
<td>12</td>
</tr>
<tr>
<td>Peer Support</td>
<td>Outside the classroom peer support and tutoring.</td>
<td>“Peer practicing since I have a study buddy.”</td>
<td>6</td>
</tr>
<tr>
<td>Outside Sources</td>
<td>Any additional sources.</td>
<td>“Getting extra problems.”</td>
<td>2</td>
</tr>
<tr>
<td>Practice Problems</td>
<td>The work outside of the normal homework</td>
<td>“Practice problems helped a lot.”</td>
<td>2</td>
</tr>
<tr>
<td>Lecture Notes</td>
<td>The information provided about the lectures.</td>
<td>“The lecture notes helped.”</td>
<td>1</td>
</tr>
<tr>
<td>Student Instructor (SI)</td>
<td>The student instructor provided as part of the class.</td>
<td>“Prefer to have someone like the SI working with me.”</td>
<td>4</td>
</tr>
<tr>
<td>Test Structure</td>
<td>The format and timing of the test.</td>
<td>“I am glad the exam was broken up into two.”</td>
<td>1</td>
</tr>
<tr>
<td>Video Lectures</td>
<td>The videos delivered to all of the classes.</td>
<td>“Watching the videos then activities were helpful.”</td>
<td>18</td>
</tr>
<tr>
<td>Teacher or Instructor</td>
<td>The instructor factored into comprehension</td>
<td>“The instructor does a good job answering questions.”</td>
<td>16</td>
</tr>
</tbody>
</table>
## Factors Affecting Comprehension

<table>
<thead>
<tr>
<th>Category / Codes</th>
<th>Description</th>
<th>In Vivo Coding</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outside Circumstances</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effort</td>
<td>There was need for additional exertion.</td>
<td>“… put more effort into my study.”</td>
<td>1</td>
</tr>
<tr>
<td>Focus on Math</td>
<td>Seeking the core understanding of math fundamentals.</td>
<td>“… context was interesting, but I was so focused on the math.”</td>
<td>2</td>
</tr>
<tr>
<td>Learning Disability</td>
<td>There was some disability mentioned in their learning.</td>
<td>“I have an issue reading things on screens”</td>
<td>4</td>
</tr>
<tr>
<td>Previous Knowledge</td>
<td>The student has taken the class previously.</td>
<td>“I have taken this class before”</td>
<td>1</td>
</tr>
<tr>
<td>Prior Class History</td>
<td>The previous classes of the student were a factor.</td>
<td>“I never took calculus so I kind of had a really rearrange my thinking.”</td>
<td>3</td>
</tr>
<tr>
<td><strong>Schedule</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family</td>
<td>There were family situations mentioned.</td>
<td>“School is depending on my kid’s schedule.”</td>
<td>4</td>
</tr>
<tr>
<td>Sickness / COVID-19</td>
<td>There were issues associated with COVID-19 from a health perspective.</td>
<td>“I actually had COVID for 1½ weeks so I wasn't able to come class.”</td>
<td>1</td>
</tr>
<tr>
<td>Time</td>
<td>Time being a factor as it relates to attending different instructional elements.</td>
<td>“A lot of things get in the way of me trying to get the time.”</td>
<td>5</td>
</tr>
<tr>
<td>Work</td>
<td>There were work situations mentioned.</td>
<td>“With work schedule, school was put on the backburner.”</td>
<td>5</td>
</tr>
</tbody>
</table>
The Effect of Context on Comprehension

<table>
<thead>
<tr>
<th>Category / Codes</th>
<th>Description</th>
<th>In Vivo Coding</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helpful</td>
<td>There was benefit.</td>
<td>“The activities themselves, they were very helpful.”</td>
<td>4</td>
</tr>
<tr>
<td>Engaging</td>
<td>Strong buy-in to the lesson</td>
<td>“They were more engaging.”</td>
<td>4</td>
</tr>
<tr>
<td>Improved Comprehension</td>
<td>Comprehension was improved as a result.</td>
<td>“It is improved my comprehension.”</td>
<td>4</td>
</tr>
<tr>
<td>Tangible</td>
<td>The intent was clear.</td>
<td>“The paper airplane was tangible.”</td>
<td>1</td>
</tr>
<tr>
<td>Useful</td>
<td>There was a relation to a practical purpose.</td>
<td>“They are very useful.”</td>
<td>3</td>
</tr>
<tr>
<td>Interesting</td>
<td>It caught the attention.</td>
<td>“… interesting that kept me like focused.”</td>
<td>26</td>
</tr>
<tr>
<td>Cool</td>
<td>There was some enthusiasm.</td>
<td>“… learning about one of them was kind of cool.”</td>
<td>1</td>
</tr>
<tr>
<td>Fun</td>
<td>The topic was fun in nature</td>
<td>“The activities were fun.”</td>
<td>7</td>
</tr>
<tr>
<td>Hands-On</td>
<td>There was active participation</td>
<td>“They were more like hands-on.”</td>
<td>4</td>
</tr>
<tr>
<td>Memorable</td>
<td>The context made the information memorable</td>
<td>“The context is the thing that made me remember that the most”</td>
<td>4</td>
</tr>
<tr>
<td>Visualize</td>
<td>It was easy to picture</td>
<td>“I need to see it.”</td>
<td>1</td>
</tr>
<tr>
<td>Not So Much</td>
<td>The context had little effect on the comprehension</td>
<td>“I think it helped and hurt.”</td>
<td>10</td>
</tr>
<tr>
<td>A little helpful</td>
<td>It helped slightly</td>
<td>“…um it helped a little bit”</td>
<td>2</td>
</tr>
<tr>
<td>Confusing</td>
<td>The context was confusing</td>
<td>“I don't know what formula to use.”</td>
<td>3</td>
</tr>
<tr>
<td>Did not make a difference</td>
<td>There was no difference made in understanding</td>
<td>“I don't think the specific context used made much of a difference.”</td>
<td>6</td>
</tr>
<tr>
<td>Fleeting</td>
<td>There were good moments in time but didn’t translate to long term comprehension</td>
<td>“But [the context] was fleeting.”</td>
<td>1</td>
</tr>
<tr>
<td>Hard to grasp</td>
<td>It was hard to understand</td>
<td>“It was a little harder for me to understand.”</td>
<td>1</td>
</tr>
<tr>
<td>Not interesting</td>
<td>The context was not interesting</td>
<td>“It was not something that I was interested in.”</td>
<td>1</td>
</tr>
<tr>
<td>Childish / silly</td>
<td>The context was silly and/or childish</td>
<td>“Sometimes it felt like almost comical …”</td>
<td>2</td>
</tr>
</tbody>
</table>
## The Effect of Context on Comprehension

<table>
<thead>
<tr>
<th>Category / Codes</th>
<th>Description</th>
<th>In Vivo Coding</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waste of time / time drain</td>
<td>The context was a waste of time and could have been better spent</td>
<td>“It just seemed like a waste of a waste of time.”</td>
<td>2</td>
</tr>
<tr>
<td>Relatability (Real World)</td>
<td>The context was able to be related to the real world.</td>
<td>“It was more helpful to have a real-world example than something that's an abstract.”</td>
<td>8</td>
</tr>
<tr>
<td>Apply &amp; applicable</td>
<td>the context allowed them to apply the material in some sort of way</td>
<td>“I want to see when I am going to apply it.”</td>
<td>2</td>
</tr>
<tr>
<td>Career and career goals</td>
<td>a direct link between the lessons and their career</td>
<td>“I am wanting to go into a field that uses statistics so yeah this is all super relevant.”</td>
<td>5</td>
</tr>
<tr>
<td>Future</td>
<td>See the lessons in their future</td>
<td>“It would help if the context was related to my future.”</td>
<td>1</td>
</tr>
<tr>
<td>major</td>
<td>An interlink between the context and their courses as part of their major</td>
<td>“If there is a relation to my major, I would of course be like super interested.”</td>
<td>2</td>
</tr>
<tr>
<td>real-world and real-life</td>
<td>The context connected to their real life</td>
<td>“More helpful to have real world examples”</td>
<td>1</td>
</tr>
<tr>
<td>relatable</td>
<td>Made it easier to follow along</td>
<td>“… that is more relevant to me than something like march madness.”</td>
<td>6</td>
</tr>
<tr>
<td>retain</td>
<td>the context helped them retain the information</td>
<td>“If it can hold my focus, I hear more, comprehend more and I retain more.”</td>
<td>2</td>
</tr>
<tr>
<td>Factors Affecting Attitude</td>
<td>Description</td>
<td>In Vivo Coding</td>
<td>Count</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td><strong>Elements of Instructional Process</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class Structure</td>
<td>Way in which instructor delivers material.</td>
<td>“The teacher makes class so much more enjoyable.”</td>
<td>7</td>
</tr>
<tr>
<td><strong>Content Got Hard</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formulas</td>
<td>The material format having an effect.</td>
<td>“I get confused with all the different formulas.”</td>
<td>1</td>
</tr>
<tr>
<td>Material Hard</td>
<td>As the context got harder, attitude began to change</td>
<td>“It's not purely positive because it's hard.”</td>
<td>5</td>
</tr>
<tr>
<td>Overload</td>
<td>The amount of information was too much.</td>
<td>“In a week you want me to learn this concept and then learn a new concept”</td>
<td>3</td>
</tr>
<tr>
<td><strong>Flipped Classroom</strong></td>
<td>Lectures at home and in-class activities.</td>
<td>“I just don't get much out of a flipped classroom.”</td>
<td>2</td>
</tr>
<tr>
<td>In-person vs online</td>
<td>Classroom environment in which students were having the class</td>
<td>“It could have been more negative if I had to take the class strictly online.”</td>
<td>1</td>
</tr>
<tr>
<td>Group Work</td>
<td>The effect of the group on their attitude</td>
<td>“We were able to work in groups which helped.”</td>
<td>1</td>
</tr>
<tr>
<td>In Person Classes</td>
<td>Class activities in person</td>
<td>“In person, sitting in a classroom is helpful.”</td>
<td>6</td>
</tr>
<tr>
<td>Proctoring System</td>
<td>The online portions made things feel impersonable</td>
<td>“It doesn't help me since there is no feedback.”</td>
<td>1</td>
</tr>
<tr>
<td>Video Lectures</td>
<td>The instruction lectures being provided by video.</td>
<td>“Good attitude would be the video lectures ...”</td>
<td>6</td>
</tr>
<tr>
<td>Instructor</td>
<td>The teacher of the in-person classes and activities.</td>
<td>“The instructor walking around seeing if we needed anything helped.”</td>
<td>9</td>
</tr>
<tr>
<td>Instructor Ability</td>
<td>The knowledge of the instructor</td>
<td>“Complain about how he does it wrong ...”</td>
<td>2</td>
</tr>
<tr>
<td>Instructor Attitude</td>
<td>Teacher attitude affected attitude of students</td>
<td>“Always had a neutral to positive attitude in class.”</td>
<td>2</td>
</tr>
<tr>
<td>Instructor Feedback</td>
<td>Any feedback provided by the instructor</td>
<td>“…didn't make me feel like stupid or anything…”</td>
<td>4</td>
</tr>
<tr>
<td>Instructor Mantra</td>
<td>Instructors foretelling of hard unit and exam affected student attitude</td>
<td>“…kept saying it would be the hardest exam and I did better than I thought…”</td>
<td>8</td>
</tr>
<tr>
<td>Instructor Willingness</td>
<td>Instructors being willing to answer questions</td>
<td>“…willingness to stay after class a bit helped.”</td>
<td>5</td>
</tr>
<tr>
<td>Personal Connection</td>
<td>Feeling like the instructor is or is not accessible.</td>
<td>“My personal interaction with the instructor affected me the most.”</td>
<td>2</td>
</tr>
</tbody>
</table>
## Factors Affecting Attitude

<table>
<thead>
<tr>
<th>Category / Codes</th>
<th>Description</th>
<th>In Vivo Coding</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outside Circumstances</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exam Performance</td>
<td>Attitude based on their exam score.</td>
<td>“If you had interviewed me before [the exam] I would have had a more negative response.”</td>
<td>26</td>
</tr>
<tr>
<td>Comprehension</td>
<td>The feeling like you knew and realizing you did not.</td>
<td>“What effected my attitude is feeling like I had it and then realizing I did not.”</td>
<td>1</td>
</tr>
<tr>
<td>Family &amp; Personal Life</td>
<td>Anything dealing with family or other personal life issues</td>
<td>“Just being able to get the stuff done, my personal life gets in the way.”</td>
<td>7</td>
</tr>
<tr>
<td>Time</td>
<td>Not having enough time</td>
<td>“I felt like I can't keep up with anything.”</td>
<td>3</td>
</tr>
<tr>
<td>Group Work</td>
<td>The effect of the group on their attitude</td>
<td>“What affected my attitude more was that most of those in class activities were as a group.”</td>
<td>6</td>
</tr>
<tr>
<td>Peers</td>
<td>Other students taking the same class</td>
<td>“I wouldn't really consult my instructor so it's definitely the peer support that helped.”</td>
<td>8</td>
</tr>
<tr>
<td>Preconceived Notions</td>
<td>Any ideas or thoughts from outside perspectives</td>
<td>“I have had a lot of people tell me the statistics was hard.”</td>
<td>5</td>
</tr>
<tr>
<td>Not a Math Person</td>
<td>Them sharing that they are not much into math</td>
<td>“I am not really a math person to be honest.”</td>
<td>8</td>
</tr>
<tr>
<td>Previous Experiences</td>
<td>Previous statistics class experience influenced their attitude</td>
<td>“Statistics is not really my thing and I’ve always been not great at it. It has always been hard for me.”</td>
<td>12</td>
</tr>
<tr>
<td>Prior Attitude</td>
<td>Feelings/Attitude about the course or statistics prior to the unit</td>
<td>“If I go into it, with at least a neutral attitude I feel like I will do a bit better.”</td>
<td>5</td>
</tr>
<tr>
<td>The want to succeed</td>
<td>The attitude of wanting to succeed no matter the circumstances</td>
<td>“I want to say it is positive in that I would love to understand statistics.”</td>
<td>4</td>
</tr>
</tbody>
</table>
## The Effect of Context on Attitude

<table>
<thead>
<tr>
<th>Category / Codes</th>
<th>Description</th>
<th>In Vivo Coding</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interesting</td>
<td>Sparks interest beyond what is encountered in a typical lesson</td>
<td>“… gave me some interesting things to think about while doing the problem.”</td>
<td>17</td>
</tr>
<tr>
<td>Engaging</td>
<td>Stronger buy-in to the lesson</td>
<td>“It made it interesting and it keeps me engaged with the material.”</td>
<td>4</td>
</tr>
<tr>
<td>Excitement</td>
<td>Provides personal satisfaction</td>
<td>“It was something that I enjoyed and was excited to do…”</td>
<td>2</td>
</tr>
<tr>
<td>Fun</td>
<td>It made things enjoyable.</td>
<td>‘It was like a fun one…”</td>
<td>2</td>
</tr>
<tr>
<td>Into It</td>
<td>The student was enrolled.</td>
<td>“I probably wouldn’t be as into it.”</td>
<td>2</td>
</tr>
<tr>
<td>Keeps attention</td>
<td>It kept the focus of the student</td>
<td>“I can say that it keeps my attention.”</td>
<td>2</td>
</tr>
<tr>
<td>No Effect</td>
<td>The context has no effect on their attitude</td>
<td>“I do not think it overall change my attitude.”</td>
<td>5</td>
</tr>
<tr>
<td>Relatability</td>
<td>Seeing how it could be used in a realistic scenario</td>
<td>“Relatable is easier to follow along because it was more engaging.”</td>
<td>13</td>
</tr>
<tr>
<td>Application</td>
<td>There was potential in the use outside of class.</td>
<td>“In class activities kind of helped me figure out where I would apply it.”</td>
<td>2</td>
</tr>
<tr>
<td>Career</td>
<td>Context relating to the career</td>
<td>“In like a business context and it's easier for me to be like this is how I would directly apply to what I am planning to be doing.”</td>
<td>13</td>
</tr>
<tr>
<td>Practical</td>
<td>Seen as useful in answering problems that exists in the real world</td>
<td>“…helped me see how I could use statistics in a more practical plan.”</td>
<td>1</td>
</tr>
<tr>
<td>Useful</td>
<td>Use of statistics material in everyday occurrences and instances</td>
<td>“It's very useful. It's something that I'm going to be using, especially on the very basic levels.”</td>
<td>14</td>
</tr>
<tr>
<td>Helpful</td>
<td>The use of context helped with attitude.</td>
<td>“It was helpful, I think it definitely affected my attitude…”</td>
<td>5</td>
</tr>
<tr>
<td>Positive</td>
<td>The experience was positive.</td>
<td>“…better experience in statistics class.”</td>
<td>1</td>
</tr>
</tbody>
</table>
APPENDIX F.1

Selected Researcher Comments

Personal Reflections from the Instructor-Researcher on the Research Methods

During the intervention, a journal (see Appendix F.2) was kept with my personal reflections about the intervention, what went well, and what could have gone better. After reviewing my notes at the conclusion of the study, there are a few items that I believe deserve to be mentioned.

It would have been nice to have had the time to pilot each of the activities first, before the experimental phase. A couple of the activities were unexpectedly long and challenging to get through in the single class period. This could be where being a veteran teacher had a larger impact on the lesson’s delivery. Having taught this course for 17 years, I tend to be mindful of the clock and I know how to adjust the pacing of the class by providing enough direction to help students move through some of the more difficult concepts or questions. Less experienced instructors tend to allow the students to continue working at their own natural pace and then often will scramble at the end of the class in order to finish. Both of the other instructors mentioned that this happened to them at least once during the longer lessons. In fact, one of the fellow instructors took part of the next class to finish one of the intervention activities. If I were to run this intervention again, I would adjust the length and scope of a few of the activities to be better suited for a 50-minute class period.

Upon reviewing the exit slips after each lesson, it was surprising to see that what we expected to be generic, non-relevant contexts were still considered “interesting” to some students. Since the majority of students enrolled in this course do not have children,
I would not have expected baby birth weights and dice to be particularly interesting or hold much personal relevance. Additionally, I also would not have thought that students would find World War II airplanes to be personally relevant, as none of them were alive during World War II and none of them were majoring in history. However, for both of these lessons, the average relevance score was greater than 3 on a 5-point scale. Selecting contexts that were not considered relevant was more difficult than I had anticipated. Similarly, selecting contexts that students thought were highly relevant was also harder than anticipated, even with the interest inventory. We saw that just because students liked the category on the interest inventory, it did not always translate into them liking the specific example we selected from that category. If I were to run this study again, I think it would be wise to have students rate specific contexts that could be used in the lessons, some that were thought to be relevant and others that were thought to be generic. Then I could design lessons based on the specifics rather than a broad category or sub-category. The interest inventory definitely helped narrow our focus, but I think an even narrower focus would be helpful in the future.

A third thought of mine had to do with our non-significant comprehension results. When students take a test, there are factors beyond what is done in the classroom that affect their test scores. No measurement was made to determine if study habits factored into the comprehension results. It is entirely possible that students in the treatment group did learn and understand the material better than their counterparts in the control group while in class, but this difference wasn’t seen due to the study habits of the students outside of class. We have no way of knowing if the students in the control group needed to study longer to achieve similar scores to their treatment group counterparts. We did not
measure how many tutoring sessions they attended or how often they sought outside help.

We just assumed that the treatment and control groups were similar in this regard, and that could have been an oversight on our part. If I were to run this study again, I would like to include a question on the post-version of the SATS-36 to try to ascertain if there was any sort of difference in out-of-class work between sections.
APPENDIX F.2

Selected Researcher Journal Entries

January 19 –
In class today, I handed out the IRB forms to my students. I also emailed the other instructors to make sure that they had the IRB form ready to go for their sections. Most students signed it, but a couple were not in class, and a couple did not turn it in. I also described the research to the class and asked them to fill out the interest survey and the comprehension pretest. I attended the control class too to help get everything set up well and to explain the research. I collected a large portion of their IRB forms too. Research is underway!

January 21 –
I was able to collect two more IRB documents from students in my class today.

January 24 –
I double checked the students who hadn’t yet completed IRB, interest survey, and pretest. I reminded them that it needed to be done. I heard the final no from a student on signing the IRB form in my class.

February 6 -
I pulled the data for the $SATS-36$ pretest and the comprehension pretest for the control group off Canvas. I assigned each student a unique identifier using the following method:

- Copied student names into Excel
- Used $=$RAND() to create a decimal random number next to each student
- Sorted the names based on random numbers from smallest to largest
- Gave each student a number in the hundreds. Hundreds digit represents the section, remaining digits go from 01 to n and match the names in the randomly sorted order.
- Remove data from each file for students who refused to sign the IRB document.

February 7 –
I sent an email reminder to students to complete their pretest (comprehension and/or $SATS-36$) before the evening class tonight. All students in sections 1 and 2 had these items done. I also emailed the Section 2 instructor to make sure one of the two students without an IRB had indicated to him that he didn’t plan on signing it, rather than be a late-registering student and not have had the opportunity to sign. Turns out he didn’t want to. I emailed all 4 students in the evening class without IRB forms to ask them if it was their intention not to sign it or if they forgot to turn it in. I pulled
the data from Section 1 for the comprehension and attitude pretests. I assigned unique identifiers to these students and deleted identifying information. I also made a few edits for readability and correctness. For instance, when asked how many math classes they took, if a student wrote “3-4”, Excel read this as March 4th. I edited it to look like “3 to 4”. Another student had written a correct answer in a non-expected form that would normally be counted as correct on a test upon instructor review, so I awarded that student the point on the pretest. Finally, I added responses of a student to the Section 1 comprehension pretest who actually took the pretest when he was enrolled in a different section. Prior to his section change in Canvas, I pulled his answers and wrote them by hand on my attendance sheet so that I could enter them for him later on.

February 8 –
I pulled the SATS-36 data from Section 3, cleaned it up, and uploaded it to our Google Drive. As of last night, two students still hadn’t taken it who had signed the IRB. I was a little worried, because five students in that class didn’t sign an IRB form and losing another two brings the sample size for the evening section down to 18. But it doesn’t make sense to look at their attitude after the intervention started, so we’ll just have a smaller group than we thought.

February 9 –
Taught Activity 1. I think the students had fun. I enjoyed talking about the lesson with them. The exit slips were mostly all positive. Then, when I gathered them from the generic class, I was surprised that students thought dice were more personally relevant than Jimmy Fallon’s egg roulette. They weren’t far off, just not what I expected. Both other instructors thought the lesson went well too.

February 11 –
Taught Activity 2. It went well. I was able to point out common errors to students as they were making them and highlight important vocabulary words. I entered all the exit slips into our spreadsheet and had an interesting conversation with the Section 2 instructor about the results of them. Again, the generic lesson was rated more relevant than the non-generic one. He thinks that maybe students are mistaking content and context or are rating them based on activity level instead of relevance. We also noticed that students are sometimes writing the content in the context line of the exit slip.

February 14 –
Taught Activity 3. It went well again. The pedagogy was good…we had good examples highlighting common errors and I was able to correct them in class. I entered exit slips for my class in the spreadsheet
February 15 –
   Picked up Exit slips for the control group and entered them in spreadsheet

February 16 –
   I did Activity 4 with the students in my class. Again, no major issues. The students
told me I should have brought my dog to class, since the context was dogs and other
pets. I showed them a picture of her instead. I observed Section 2’s class (control)
and explained to them the difference between content and context, since that seems to
be an issue of confusion. I also suggested to Section 3 instructor that he do the same
in his class for consistency in the Exit slips.

February 21 –
   Taught Activity 5. It was a little long, and I had to do a lot of the work to get the
students finished by the end of the class period. The control group did not finish and
will be using some of Wednesday’s class to complete theirs. The evening class did
finish (but they had 75 minutes, not 50.) I think all-in-all, the students liked the idea
of batch testing for COVID, and I could tell that they did understand the idea of why
batch testing is more economical in some situations than in others. They struggled a
little with the basic probability computations. (These were concepts from two weeks
ago.) But once the probability was in place, they were able to handle the new
concepts of expected value, variance, and standard deviation.

February 23 –
   Taught Activity 6 (Binomial). This one would have gone a little better if we had
more time in class and if I had been able to collect the data from each student a day or
two earlier. Some students didn’t quite follow the directions, and there wasn’t time
for me to edit or give feedback to them before starting the lesson. There were also a
few duplicate probabilities, so it wasn’t entirely “all guesses” for every student. The
premise of the activity was great, but a trial run through would have helped work out
some of the kinks. Because of the snow, we had a really small turnout for the
Binomial (2 Truths and a Lie) evening class...Just 5 of them came for that. I had a
small, but reasonable class, and the control group was also reasonable.

February 28 –
   Taught Poisson distribution (was a little long). I wasn’t sure how many “errors”
students would make in their food discussion, but it worked out to be a very
reasonable number. I was really surprised and happy with that! The average number
of t-words in 1 minute ranged between 2.667 and 9. So choosing a value of 6 for the
Poisson probability was a great choice! The numbers all worked out really well for that activity!

March 2 –
Taught Uniform distribution (was a little long) and a little confusing toward the end. In retrospect, should have had more straightforward questions.

March 4 –
Entered Exit slips from other classes for the week.

March 7 –
Taught Activity 9 (Normal). This went fine. There wasn’t anything super remarkable about the lesson.

March 9 –
Taught Activity 10 (exponential). The students seemed to be okay getting up and shooting baskets. I think many of them liked getting out of their seats. I heard from one of the other instructors a similar comment about the students enjoying getting up and moving.

March 11 –
Sent a fellow researcher a list of students’ names who should/should not be considered for interviews based on the number of lessons attended. I made sure not to tie the students to the exit slip information, so he couldn’t determine any of the identifiers for any student. I listed students as “yes” for 8-10 activities, “ok” for 7 activities, “rather not” for 5-6 activities, and “no” for 4 or fewer activities or for students who did not give IRB permission.

March 18 –
I reviewed the exams for my in-person section, and sent an email to the other instructors with instructions for how to consistently award points for questions the computer might not have graded correctly.

March 28 –
I downloaded all exam results from Canvas so that I could analyze them in the next week or so.

April 12 –
I sorted through all the quantitative data and made spreadsheets comparing pre-post for the 10 questions and looking at the overall Exam 2 scores for all 3 sections and
uploaded them to the Google drive. Interestingly, the test groups out-performed the control group, but I’m not sure it’s going to be statistically significant.

April 18 –
Discussed as a research team at what point we wanted to break the participants up for participation/non-participation. We decided students who were present for 6-10 should count, while those who were present for 5 or less should not. I will go back and re-analyze the results for students based on participation.