UNDERGRADUATE STUDENT DIFFICULTIES, CONVICTIONS, AND SELF-EFFICACY WITH STRONG AND WEAK MATHEMATICAL INDUCTION

A Thesis

by

BRIELLE DUCHESNEAU

BS, Texas A&M University-Corpus Christi, 2019

Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in

MATHEMATICS

Texas A&M University-Corpus Christi
Corpus Christi, Texas

December 2022

UNDERGRADUATE STUDENT DIFFICULTIES, CONVICTIONS, AND SELF-EFFICACY WITH STRONG AND WEAK MATHEMATICAL INDUCTION

A Thesis

by

BRIELLE DUCHESNEAU

This thesis meets the standards for scope and quality of
Texas A&M University-Corpus Christi and is hereby approved.

George Tintera, PhD
Chair

Valentina Postelnicu, PhD
Co-Chair/Committee Member

James Dogbey, PhD
Committee Member

December 2022

ABSTRACT

The purpose of this study is to describe and compare undergraduate students' difficulties with proofs by strong and weak mathematical induction and to investigate their convictions in their proofs. Further, the study looks at self-efficacy for mathematical proofs using the two forms of mathematical induction, weak and strong induction. Participants were five undergraduate mathematics major students from a four-year university. Students completed a demographic questionnaire, a self-efficacy questionnaire for proofs by mathematical induction, and three tasks asking them to prove a statement $P(n)$ using strong and weak mathematical induction and to state their convictions in the value of truth of the statement $P(n)$ for different values of n. Two of the five students participated in an interview to probe for difficulties with proofs by induction.

The results showed that students had more difficulties with strong induction. Students from the study also had the most difficulties with the inductive step itself for both weak and strong induction. Students' convictions regarding the validity of a statement $P(n)$ seemed to be the same overall for weak and strong induction. Students had proof conviction when asked if $P(n)$ was true for a value n within the domain of validity. They had mixed answers when they were asked for convictions referring to a value n outside of the domain.

Scores for self-efficacy and performance for strong induction tasks were lower than the self-efficacy and performance scores for weak induction tasks.

ACKNOWLEDGEMENTS

I would like to thank my thesis committee members, George Tintera, Valentina Postelnicu, and James Dogbey, for helping me write my thesis to earn my master’s degree from TAMUCC.

I would like to thank George Tintera for encouraging me and meeting with me numerous times. He also helped conduct the interviews.

A special thank you to Mario Gonzalez for giving permission to expand his research. He was a role model in forming my own thesis.

A special thank you to Valentina Postelnicu, who was chair for Mario Gonzalez's thesis for helping me begin my thesis journey.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Rationale

This thesis seeks to contribute to closing the gap in the literature on mathematical induction and the difficulties undergraduate mathematics students have with it. Studies such as those conducted by Avital and Hansen (1976), Baker (2010), Ernest (1984), and Harel (2001) pointed out where gaps appear with weak induction and student difficulties at any grade level. Little research can be found on student difficulties with strong induction. Ernest (1984) has shown some apparent struggles that undergraduate students have with the weak form of mathematical induction, and studies link their performance to exposure to mathematics courses. However, no studies focus on the difficulties undergraduate mathematics majors encounter with the strong form of induction.

This study also investigated students 'convictions about mathematical induction and whether students have the same convictions for both weak and strong induction. If not, how do they differ? Researchers such as Weber and Mejia-Ramos (2015) investigated what convictions students have about mathematical proofs in general. Gonzalez's recent study for his thesis (2020) considered student conviction in mathematical induction proofs. However, no studies have yet to compare convictions held for strong and weak induction.

Research conducted by Bandura (1977) has shown that student self-efficacy plays a vital role in student learning outcomes in mathematics. However, no known studies focus on self-efficacy, especially for mathematical induction. Given the growth of undergraduate mathematics, it would be beneficial to find the connections between self-efficacy and its importance in proving statements using strong and weak mathematical induction.

Statement of the Problem

During the last decade, research on student difficulties with proofs by mathematical induction was scarce. Most of the studies focused on the teaching and learning of proofs by weak mathematical induction and student difficulties with weak induction. Although there is no literature found on strong induction, it is still taught at the undergraduate level. Finding the comparison between strong mathematical induction and weak induction performance is not reported. If we can understand these difficulties, we can target the way mathematical induction is taught both in the weak and strong forms, when asked to prove a statement P(n), n being a natural number, $n \geq n_0$.

Studies of student convictions about mathematical induction are scarce as well. Gonzalez (2020) focuses on convictions with weak induction. Stylianides and Stylianides (2007) are unique in that they studied convictions for undergraduate preservice math teachers. There are no studies on student convictions for strong induction or the comparison to weak induction.

Also, research on mathematics self-efficacy and its possible connection to student performance at the college level is scarce as well. Studies tend to focus on students that are at pre-college levels. There are no studies connecting self-efficacy to student performance of mathematical induction.

Purpose of the Study

This study aims to describe undergraduate student difficulties with proofs by strong and weak mathematical induction and to investigate whether there are any differences between the difficulties with proofs by weak and strong induction. The study will also describe if there are differences in student conviction for strong versus weak induction.

Further, the study examines self-efficacy for mathematical proofs and the two forms of mathematical induction, strong and weak mathematical induction.

Research Questions

This study addresses the following research questions:

1) What difficulties do undergraduate mathematics students have doing proofs when using the method of weak induction versus strong induction?

2) How do students' convictions differ for weak versus strong induction?

3) What factors or attributes do students have that relate to having difficulties with weak or strong induction? How is positive self-efficacy in using weak and strong induction connected to student performance?

Significance of the Study

This study is significant because it showed that students had more difficulties with strong induction over weak induction, which hasn't been looked at before. New to the literature, it was also found that students have a similar conviction for strong induction and weak induction. Students have proof conviction for values within the domain of validity but mixed answers when given a value outside of the domain of validity for both strong and weak induction. Another new finding is that the study showed that when students had lower self-efficacy in both forms of mathematical induction, then their performance was also lower for both forms. Self-efficacy and performance were higher for weak mathematical induction than for strong mathematical induction. This shows that there may be some connection between self-efficacy and student performance of higher-level math.

Theoretical Frameworks

Framework to Account for Student Difficulties with Proofs by Induction

Gonzalez (2020) used a framework to account for student difficulties in mathematical induction proofs, based on Ernest (1984) framework for a proof by mathematical induction, given a statement P(*n*), *n* being a natural number, $n \geq n_0$:

- in the **basis step**, the proof of $P(n_0)$ using algebraic substitutions is needed;
- in the **inductive step**, the derivation of $P(k+1)$ from the inductive hypothesis $P(k)$ is required, $P(k) \rightarrow P(k+1)$;
- finally, **invocation of** the Principle of Mathematical Induction (**PMI**) comes from successfully verifying the basis step and the inductive step to show that $P(n)$ is true for all *n* natural, $n \geq n_0$.

The domain of validity for $P(n)$ is the set of values of *n,* $n \geq n_0$ for which $P(n)$ is true.

Using the behavioral skills needed to complete a mathematical induction proof in the correct form, as stated by Ernest (1984), Gonzalez (2020) developed an adaptation of the Ernest's (1984) framework to be able to identify student difficulties with mathematical induction. Gonzalez (2020) found that the **basis step** was associated with the following difficulties: considering the basis step; correctly evaluating $P(n_0)$ and choosing the correct starting point n_0. The following difficulties were associated with the **inductive step**: algebraic manipulations to show $P(k) \rightarrow P(k + 1)$, the substitution of $P(k)$ into $P(k + 1)$ to show $P(k) \rightarrow P(k + 1)$, stating $P(k + 1)$, and stating $P(k)$. Finally, the difficulties that can occur when writing the proof in the correct form are the following; explicitly writing $P(k + 1)$ is to be shown assuming that $P(k)$ is true, stating the domain for the introduced variable *k*, and **invoking PMI**.

The framework in Gonzalez' (2020) study focused on using weak induction mathematical proofs. To analyze student difficulties with strong induction, the framework was modified as follows. In the inductive step, difficulties may be substitution of $P(n_0),\ldots,P(k)$ into $P(k+1)$ to show $P(n_0),...,P(k) \rightarrow P(k + 1)$ and another difficulty is the actual algebraic manipulations that show $P(n_0),...,P(k) \rightarrow P(k+ 1)$. The difficulty for writing the proof in the correct form for strong induction refers to explicitly writing $P(k + 1)$ is to be shown assuming that $P(n_0),...,P(k)$ are true. Gonzalez (2020) adopted after Ernest's (1984) framework is called the E-a framework for short. In this study, the same framework is used, with modifications for strong induction.

An example of a correct student's response using Ernest's (1984) analysis may be applied to a task such as the "Breaking a chocolate bar" problem posed by Bogomolny and Taleb (2020).

Let $P(k)$ be the statement that a chocolate bar of size k requires $k - 1$ breaks where $k \geq 1$.

1. **Basis Step**: $k = 1$: $P(1)$ is true because there is only a single square of chocolate, and $1 - 1 = 0$ breaks are required.
2. **Inductive Step**: We suppose $k \geq 1$ and any chocolate bar of size s, where $1 \leq s \leq k$, requires $s - 1$ breaks. We must now show there is a way to break a chocolate bar of size $k + 1$ with k breaks.

   To do this, first break the chocolate bar of size k+1 into two smaller pieces of size p and q where $p + q = k + 1$. This is certainly possible because the size of the bar is at least two. Now the pieces of sizes p and q are between 1and k, so by strong induction, breaking these two pieces into single squares requires only $p - 1$ and $q - 1$ breaks, respectively. The total number of breaks required to break the bar of size k+1 into single squares is therefore $1 + (p - 1) + (q - 1) = p + q - 1= (k + 1) - 1 = k$.

3. **Invocation of PMI**: Since $P(1)$ is true and for all k, and $P(1),\ldots,P(k)$ implies $P(k+1)$, by PMI, this shows that $P(n)$ is true for all $n \geq 1$.

Below are examples of student difficulties with the "Chocolate Bar Problem" using the E-a framework.

Consider the task: The proof is by strong induction on the size of the chocolate bar. Let $P(k)$ be the proposition that a chocolate bar of size k requires $k - 1$ breaks where $k \geq 1$.

- Difficulties with the basis step.
  - Omitting the basis step. For example, a student's proof may look like this:

    The student will have the inductive step and the invocation of PMI but no basis step. So "$k = 1$: $P(1)$ is true because there is only a single square of chocolate, and $1 - 1 = 0$ breaks are required." is missing from the proof.

  o Choosing the incorrect n_0. For example, a student may choose 0 instead of 1 which will result in:

    $k = 0$: So, the number of breaks in the chocolate bar is 0-1=-1. But you can't have negative breaks on a chocolate bar!

  o Incorrectly evaluating in the basis step. For example, a student may substitute 1 into $P(k)$ but evaluate incorrectly similar to:

    $P(1)$: So, the number of breaks is 1-1=-1.

    However, you can't have negative breaks!

  The student might then discontinue the proof or claim the statement, $P(k)$, to be false because $P(1)$ was considered false.

- Difficulties with proving the implication $P(1),...,P(k) \rightarrow P(k+1)$
  - o Difficulty stating $P(1),...,P(k)$. For example, a student might do the following:

    The student may forget to add the initial break in the algebraic part of the inductive step.
  - o Difficulty stating $P(k+1)$.

    For example, a student attempting to state $P(k+1)$ may write: $k+1=k$ (confusing the number of pieces with the number of breaks).
  - o Difficulty manipulating algebraically the terms in $P(k)$ and $P(k+1)$ to show $P(k) \rightarrow P(k+1)$.

    Confusing the number of pieces with the number of breaks.
  - o Difficulty substituting $P(1),...,P(k)$ into $P(k+1)$ to show $P(1),...,P(k) \rightarrow P(k+1)$. For example, a student might write:

    Assume a bar with k or fewer tabs can be broken into individual tabs with k-1 breaks. Given a bar with k+1 tabs, break it into two bars with j, and m tabs, respectively, with $j+m=k+1$ and $j, m < k+1$. Then $k+1-1 = 1 + j-1 + m-1 = j+m$, the sum of the initial break and the breaks for the two smaller bars.
- Difficulties with writing the proof in the correct form.
  - o Not invoking PMI. For example, a student may do the following:

    Student will write the basis step and inductive step, but they will leave out "Therefore, by PMI, this shows that P(n) is true for all n $\geq$1."

- Not explicitly writing that $P(k + 1)$ is to be shown assuming that $P(1)$,...,P(k) are true. For example, a student may write for the inductive step:

  To do this, first, break the chocolate bar of size k+1 into two smaller pieces of size p and q where $p + q = k + 1$. This is certainly possible because the size of the bar is at least two. Now the pieces of sizes p and q are between 1 and k, so by strong induction, breaking these two pieces into single squares requires only $p - 1$ and $q - 1$ breaks respectively. The total number of breaks required to break the bar of size k+1 into single squares is therefore $1 + (p - 1) + (q - 1) = p + q$ - 1= (k + 1) - 1 = k.

- Not stating the domain for the introduced variables k, p, q. For example, a student might write for the inductive step:

  We suppose any chocolate bar of size s, requires $s - 1$ breaks. We must now show there is a way to break a chocolate bar of size $k + 1$ with k breaks. To do this, first, break the chocolate bar of size k+1 into two smaller pieces of size p and q where $p + q = k + 1$. This is certainly possible because the size of the bar is at least two. Now the pieces of sizes p and q are between 1 and k, so by strong induction, breaking these two pieces into single squares requires only $p - 1$ and $q - 1$ breaks, respectively. The total number of breaks required to break the bar of size k+1 into single squares is therefore $1 + (p - 1) + (q - 1) = p + q$ - 1= (k + 1) - 1 = k.

Framework to Account for Student Convictions with Proofs by Induction

Gonzalez (2020) developed a framework for student convictions based on a study conducted by Weber and Mejia-Ramos (2015). According to Weber and Mejia-Ramos (2015), students can have two types of conviction of mathematical proofs. They can have absolute conviction or relative conviction. Absolute conviction is when a student has no doubt about a claim and its validity and is completely convinced. Relative conviction is when a student is not entirely convinced and requires additional actions until they have reached conviction in the value of truth of a statement. According to Weber and Mejia-Ramos (2015), if a student does not have a response that fits into one of the two categories of conviction, there is no other category that exists for their responses. Thus, to analyze all possible responses, Gonzalez (2020) added two more categories of conviction, and called this enriched framework Weber and Mejia-Ramos, or WMRe. The four categories of the WMRe framework are: **proof conviction**, **absolute conviction**, **unconvinced disbelief**, and **empirical disbelief**. **Proof conviction**, redefined from Weber and Mejia-Ramos (2015) by Gonzalez (2020), states that a student derives his or her conviction from his or her proof of the understood statement. **Absolute conviction**, redefined by Gonzalez (2020), states that a student derives their conviction from the proof of the understood statement and is convinced that for a value not within the domain of validity, the statement is false. **Unconvinced disbelief** is a refinement of relative conviction by Gonzalez (2020). It states that the proof produced by the student does not aid in establishing their conviction, and their response does not correspond with what they proved. Finally, **empirical disbelief**, defined by Gonzalez (2020), is when the student misunderstands the definitions in the statement itself and is unconvinced by their own proof.

To provide examples of how students' responses may be categorized using the WMRe framework, recall the "Breaking a chocolate bar" problem posed by Bogomolny and Taleb (2020).

Let P(k) be the proposition that a chocolate bar of size k requires $k - 1$ breaks where $k \geq 1$.

The following are the categories in the WMRe framework, accompanied by examples of student responses from the task previously mentioned:

- **Proof Conviction** – For example, a student asked whether the statement, $P(k)$, was true for $k = 2$ may respond by:

  My proof has shown $P(k)$ is true for all integers greater than or equal to 1, and 2 is an integer greater than 1, therefore the statement $P(2)$ is true.

- **Absolute Conviction** – For example, a student asked whether the statement was true for $k = 0$ would respond by saying:

  The proof was only done for integers greater than or equal to 1, and 0 is not greater than 1, therefore, $P(0)$ is not true.

- **Unconvinced Disbelief** –For example, a student asked whether the statement, $P(k)$, was true for $k = 2$ would respond by saying:

  If k=2

  $P(2)$: then there are 2-1=1 breaks.

  Since 2 pieces means 1 break, $P(2)$ would therefore be true.

- **Empirical Disbelief** – For example, a student asked whether the statement was true for $n = 2$ would respond by saying:

  If k=2

P(2): then there are 2-1=1 breaks. But there could also be 2 breaks in the chocolate bar which is not *k*-1 break, therefore *P*(2) is not true.

Framework to Account for Student Self-Efficacy

Various researchers have measured self-efficacy for mathematics, using the Likert scale, and compared it to mathematical achievement scores (Liu & Koirala, 2009; Lent, Brown, Larkin, 1986; Schunk, 1989; Usher & Pajares, 2009).

Liu and Koirala (2009) examined the relations between mathematics self-efficacy and mathematics achievement using a four-point Likert scale to measure self-efficacy and a math Item Response Theory score to measure mathematics performance for various topics. The research was conducted on 25 middle school students from an enrollment list of 752 that agreed to participate, and results found that the relationship between mathematics self-efficacy and mathematics achievement were positively related.

Other researchers, such as Usher and Pajares (2009), have used the Likert scale to measure self-efficacy in mathematics achievement. In the study, four measures of self-efficacy were addressed, including mathematics skills. Students had to respond using a Likert scale ranging from “not confident at all” to “completely confident.” Although the study was conducted on middle school students, we can still use the same model to measure self-efficacy in mathematics for other grade levels, including college students.

In this study, 6 statements that are appropriate for mathematics majors were administered to student participants in a survey. The statements are the following:

- I am confident that I can complete a mathematical proof.
- I am confident that I can complete a mathematical proof by strong induction.
- I am confident that I can complete a mathematical proof by weak induction.

- I am confident in my ability to do a proof by mathematical induction in the correct form.
- I am confident in my ability to prove the basis of induction.
- I am confident in my ability to prove the induction step.

Students may select any of the following answer choices to the 5 statements below:

- Strongly Agree
- Agree
- Undecided
- Disagree
- Strongly Disagree

Chapter Summary

In this chapter, frameworks for students' difficulties with weak and strong induction, students' convictions of proofs, and students' self-efficacy and the association with student performance were presented. The E-a framework used in Gonzalez (2020) will be used for weak induction and the hypothesis will be modified and used for strong induction. The WMRe framework was extended in Gonzalez (2020) and will be used to analyze student convictions for strong and weak mathematical induction. Finally, a Likert scale will be used to measure self-efficacy as used by many researchers in previous studies. The research questions were introduced as well as the purpose of the study in this chapter. In the next chapter, the literature review to support the frameworks will be presented.

CHAPTER II

LITERATURE REVIEW

Student Difficulties with Strong and Weak Induction

Mathematical induction is a specific method of proof for statements that are quantified for the set of positive integers or specific subsets of that set. According to Ernest (1984), students find mathematical induction a difficult method to master, and students' lack of exposure to the induction method of proofs is an area of concern. For students to understand and successfully do a proof by mathematical induction they must know what the components are of mathematical induction, how to correctly use mathematical induction as a proof, and have a particular set of behavioral skills (Ernest 1984).

As defined by Ernest (1984), the principle of mathematical induction can be described as the following; If 1 has property P, and if any n having property P implies that $n+1$ has property P, then every n has property P. Similarly, the principle of complete or strong mathematical induction states that if 1 has property P, and if for all integers from 1 to n, having property P implies that $n+1$ has property P, then every n has property P.

For mathematical induction to be correctly used there are four main components (Ernest, 1984). The first component is the statement to be proved. The second component is the explicit invocation of the principle of mathematical induction and two subordinate proofs. The third component is the verification of a single, initial case of the statement which is known as the basis of mathematical induction. The last component is proof of the universally quantified implication statement which is the induction step. However, these behavioral skills slightly vary depending on the context and theorems to be proven.

Ernest (1984) states that there are 3 main behavioral skills that students need to successfully complete a proof by mathematical induction. These skills are the ability to prove the basis of induction, the ability to prove the induction step, and the ability to present the proof of mathematical induction in the correct form. To successfully prove the basis step, one must verify the fixed properties hold true for particular numbers. To prove the induction step, one must deduce a conclusion from a hypothesis and be able to manipulate algebraic expressions and identities. In order to successfully present a proof by mathematical induction in the correct form one must communicate the knowledge of the correct form of the proof in some way.

Many of the difficulties students have when they encounter mathematical induction, according to Ernest (1984), are because of misconceptions or conceptual difficulties that they have. There are six main sources of difficulties and errors that arise for students. The first is the ambiguity in the word "induction" itself. The inductive method is used for arriving at a conjectured generality describing a finite sequence of examples. Mathematical induction is a rigorous form of deductive proof. This ambiguity can cause students a great deal of confusion when they cannot make the distinction between the two types of induction. The second misconception students have is that mathematical induction is that they assume what they have to prove and then prove it. Basically, students assume P(n) and end up proving P(n) using complicated procedures. Temple (1957) says we can avoid this misconception by expressing induction in a two-variable form." If P(l), and if all for all k, $P(k)$ implies $P(k + 1)$, then for all n, $P(n)$." The third misconception students have is that they are not able to interpret logically complex the fourth misconception students have is that they view one of the components of induction as not essential such as the basis step. The fifth misconception arises from the exclusive use of summing finite series in induction problems The last misconception is more of a lack of understanding that induction is not selfevident or a generalization of previous elementary experiences. Many students wonder why the principle of mathematical statements is used and find the Principle of Mathematical Induction difficult to comprehend. They do not understand the manipulation of free variables in deductions or the use of quantifiers in mathematical induction.

Stylianides, Styliandes, and Philippou (2007) investigated prospective elementary and secondary school teachers and their mathematical inductive proof. Participants of the study were also asked if the proof was true for numbers within and outside of the domain of the variable. Of the participants, 25 were math majors and 70 were education majors. All were preservice teachers. Stylianides, Stylianides, and Philippou (2007) found three main difficulties that the participants had. The first difficulty was the essence of the basis step. The second difficulty was the meaning associated with the inductive step in proving the implication of $P(k)$ implies $P(k+1)$. The last difficulty was with the truth set of a sentence in a statement proved by mathematical induction to include values outside of its domain of the variable.

According to Avital and Libeskind (1978), students have conceptual, mathematical, and technical difficulties. The first conceptual problem students have is the implication of $P(k)$ to $P(k+1)$. For the student to be able to use mathematical induction the student must be able to understand that for every p that is true, when p implies q, then q must be true. The next difficulty students face conceptually, is related to the "jump" from $n=1$ to the passage form, $n=k$ to $n=k+1$. The student should be able to understand how to find a conjecture that is true not because of the patterns continuing but because they notice a sequence moving from one value to the next. Therefore, a complete proof of the result depends on proving at once the infinitely many implications.

According to Avital and Libeskind (1978) mathematical difficulties students face come from the fact that when $P(k)$ implies $P(k+1)$for all $k\geq m$ that P(n) is true for each natural number $n\geq m$, sometimes the property still holds true for a value of less than m but it is not important to the theorem. For example, if a theorem is true for $n=1$ and for all $n\geq 5$. So, this will make students think that the statement would be true for all $n\geq 1$which is false. Avital and Libeskind (1978) also found that difficulties that arise from the technical side of mathematical induction originate from the fact that a statement is written in an explicit manner. Explicit statements are considered hard for even more advanced college students.

Research conducted by Gonzalez (2020) included 78 undergraduate students at a four-year university. Three tasks were given, and interviews were held with eight participants. Students were asked to complete a task about proving identities using mathematical induction and then they were probed for their convictions about the truth of the mathematical statements they proved. Student responses were analyzed using the Ernest (1984) framework for their difficulties with mathematical induction. Gonzalez (2020) found that most students were able to complete the basis step of the induction proof. However, students had many difficulties with the inductive step and the main difficulty for them was stating $P(k+1)$. The students had the most difficulties also with writing the proof in the correct form.

Norton, Arnold, and Koshushkin (2022) investigated the persistent challenge posed by students in proof courses and implemented “quasi-induction.” That study concluded that careful use of the principle of universal generalization is needed to connect quasi-induction to mathematical induction.

Student Convictions with Proofs

According to Weber and Mejia-Ramos (2015), students can have absolute conviction or relative conviction about a mathematical statement. A student with absolute conviction has absolute certainty over a mathematical claim. A student with relative conviction expresses a high probability of a claim being reliable but not certain. Relative conviction is more common and accepting arguments are more of a matter of judgment and evaluation. Empirical arguments according to Weber and Mejia-Ramos (2015) support a mathematical statement that holds true for a subset or scope of that statement. A deductive argument on the other hand supports a mathematical statement that is a sequence of assertions that concludes with a mathematical statement where each assertion is a claim known or assumed to be true.

Weber and Meija-Ramos (2015) also did a study to analyze what convictions mathematics majors have about mathematical proofs. The study included 28 mathematics majors that have already completed a transition to proof course. The students were given ten mathematical arguments and they were asked to judge how convincing each argument was and whether the arguments constituted mathematical proofs. The primary findings of this study showed that most participants did not find the empirical arguments convincing and they found diagrammatic arguments to be convincing and valid proofs. The participants of the study found invalid deductive arguments as convincing because they failed to recognize their logical flaws. The study suggests for these students to improve the ways they process the arguments that they read.

The National Council of Teachers of Mathematics (2000) argues that all students by the time they reach 12th grade should be able to recognize reasoning and proofs as fundamentals and they should also be able to evaluate and develop mathematical arguments and proofs. Harel and

Sowder (2007) found that college-level students' difficulties with proofs originate from the fact that they are convinced by diagrammatic and empirical arguments and uncertain about what features a proof must have to be considered acceptable.

There have been studies according to Weber and Mejia-Ramos (2015) conducted on middle school and high school students as well as first-year college students and their convictions of mathematical proofs. Most of these studies have found that students are in fact convinced that empirical arguments are valid and see the format of the argument as more important than whether the argument is a proof. Weber and Mejia-Ramos (2015) are one of the few who collected on mathematics majors and what they consider to be mathematical proofs in general.

Sowder and Harel (2003) conducted a study in which they interviewed mathematics majors on how their proof schemes changed over time. The study found that mathematics majors were also convinced by empirical arguments.

To determine conviction in a statement proved by mathematical induction, Gonzalez (2020) asked participants of his study whether the statement was true for a value less than the value in the basis step, not covered by the proof, and for a value greater than the value in the basis step, covered by the proof. Their responses and justifications indicated their convictions in the statement. Only 10% of the students responded with clarity of the proofs that were produced. Most students (53%) derived conviction from empirical evidence. The percentage of students who derived conviction from the proof was 29%. This study showed that students tend to mostly derive conviction from empirical evidence rather than proof.

Student Self-Efficacy with Mathematical Induction

Bandura (1977) hypothesized that self-efficacy affects effort expenditure, persistence, and choice of activities. People with low self-efficacy tend to avoid the task as opposed to people with high self-efficacy who handle tasks eagerly. Assuming prerequisite skills are adequate, positive self-efficacy, outcome expectations, and valued outcomes all influence directions and choice of human behavior. Early self-efficacy research conducted by Bandura (1977) looked at how different experimental treatments affected individuals' self-efficacy to perform various human actions.

Brown and Inouye (1978) studied how self-efficacy influenced student achievement. They found that self-efficacy and persistence were positively correlated with each other by doing studies of college male students and solving anagrams.

Zimmerman and Ringle (1981) found that the influence of an adult model's persistence and statements of confidence raised children's self-efficacy for solving puzzles.

Hackett and Betz (1981) extended Bandura's (1977) research and studied the importance of the construct of self-efficacy in understanding the differences in educational and career choices of men and women. Their argument is that as a result of socialization experiences, women have lower expectations than men do in achievement behaviors, so they fail to live up to their capabilities and talents in career paths. They also examined how self-efficacy in educational attainment and career development is related to mathematics performance and mathematical self-efficacy. Women also are underrepresented in higher-paying jobs in science and technical fields. This is explained by their lack of preparation relative to men in mathematics.

Hackett and Betz (1981) also found that fewer female students take mathematics courses than males both in college and in high school. There are also fewer women who choose to major in mathematics. These sex differences have been a result of socialized negative attitudes and mathematics anxiety. The social learning theory perspective shows that self-efficacy expectations are an even more important factor in mathematics performance and attitudes toward math. Bandura finds that mathematics anxiety is related to math self-efficacy.

Siegel, Galassi, and Ware (1985) found a self-efficacy model superior to the mathematics anxiety model so they could predict student performance on mathematics exams. They found a significantly higher variance in the relationship between self-efficacy and mathematics exam performance such as the Scholastic Achievement Test (SAT).

Usher and Pajares (2009) studied self-efficacy and mathematics achievements and their correlation. That study shows that in general, male students tend to be more confident than female students in academic-related areas such as STEM. But these gender differences in self-efficacy are confounded by a number of factors.

Lent, Brown, and Larkin (1986) reported that men tend to enroll more in mathematics courses prior to college and so they have a greater opportunity to develop math skills and efficacy percepts.

Benson (1989) found that men reported higher mathematics self-efficacy and suggested it was because women have taken fewer mathematics courses.

Chapter Summary

This chapter included a literature review on students' difficulties with weak and strong induction, students' convictions of proofs, and students' self-efficacy and the association with student performance. There was literature on student difficulties for weak but none for strong induction. The literature also showed how convictions can be studied with weak mathematical induction. The literature finally showed how self-efficacy is linked to student performance in mathematics, but none target mathematical induction. In the next chapter, the methods of analysis will be presented.

CHAPTER III

METHODS

Research Design and Participants

The research focused on comparing difficulties that undergraduate mathematics students have with weak (ordinary) versus strong (complete) induction The study investigated students' convictions with weak and strong induction, and how students' self-efficacy is associated with their performance on weak and strong induction tasks.

Undergraduate mathematics major students were recruited from a four-year university. Students were recruited in person and by email as well. All participants completed an online instrument consisting of preliminary questions, self-efficacy questions, and three tasks. The total participants were 5 students, 3 students without being interviewed and 2 students were interviewed.

During the interview, the discussion between the two students, S1 and S2, and the researcher was recorded with observer notes. The interview was structured as a task-based interview (Goldin, 2000). The interview lasted about an hour and a half. Students were given each task, one at a time. and then probed with questions from the Oral Section of Interview page as seen in APPENDIX B for instruments.

Instruments

Students completed three sections. The first section of the instrument contained preliminary questions about mathematical induction as seen in APPENDIX A. The second section consisted of six self-efficacy questions about mathematical induction constructed by the researcher. Students were asked to indicate their agreement with each statement concerning their regular coursework. Student responses ranged from Strongly Agree to Strongly Disagree for the following statements:

- I am confident that I can complete a mathematical proof.
- I am confident that I can complete a mathematical proof by strong induction.
- I am confident that I can complete a mathematical proof by weak induction.
- I am confident in my ability to do a proof by mathematical induction in the correct form.
- I am confident in my ability to prove the basis of induction.
- I am confident in my ability to prove the induction step.

The third section of the instrument consisted of three mathematical induction tasks formed from the chocolate bar problem posed by Bogomolny and Taleb (2020). The first task was best proven using weak induction and the second and third tasks were best proven using strong induction. Each task had an additional two questions about student convictions. One of the conviction questions had a number within the given domain of validity, and one of the conviction questions had a number outside of the given domain of validity. The tasks and conviction questions were given to the students to complete in writing.

The chocolate bar problem tasks used weak induction for **Task A** and used strong induction for **Task B** and **C**. In **Task A**, a candy bar with n pieces in one row has pieces broken off one at a time. With **Task B,** instead of breaking off one piece at a time, the bar is broken into two separate, smaller bars until there are no bars left with more than one piece. For **Task C** the candy bar is broken into two parts, one with 2 pieces and one with the remaining n-2 pieces. **Part a** of each task asks for a proof by mathematical induction. **Part b** of each task asks for student convictions within the domain of validity in **Part a**. **Part c** of each task asks for student convictions outside of that domain of validity. See the tasks instrument in APPENDIX A.

Methods of Analysis

Research Question 1

1) What difficulties do undergraduate mathematics students have doing proofs when using the method of weak induction versus strong induction?

To answer Research Question 1 about weak induction, students' answers on **Task Aa** were scored using the following rubric developed by Gonzalez (2020) as seen in **Table 1**. Students' answers were scored on a scale of 0 through 10 with 0 being the lowest. The mean of these scores was reported to analyze the difficulties the students had with weak induction.

Table 1

Rubric for Task Aa, Ba, and Ca

Steps with sub-steps of a proof by mathematical induction	Points
Basis step	1
Choosing the correct starting point, n_0	0.50
Correctly evaluating $P(n_0)$	0.50
Inductive step	7
Stating $P(k)$	1
Stating $P(k + 1)$	1
Substituting $P(k)$ into $P(k + 1)$ to show $P(k) \rightarrow P(k + 1)$	1
Algebraic manipulations needed to show $P(k) \rightarrow P(k + 1)$	4
Writing the proof in the correct form	2
Invoking PMI	1
Stating the domain the introduced variable k	0.50
Explicitly writing that $P(k + 1)$ is to be shown assuming that $P(k)$ is true	0.50

Total	10

To answer Research Question 1 about strong induction, students' answers on **Task Ba** and **Task Ca** were scored using the following modified rubric, based on the one developed by Gonzalez (2000) for weak induction. Students were scored on a scale of 0 through 10 with 0 being the lowest. The means of these scores were reported using **Table 2** to analyze the difficulties the students had with strong induction.

Table 2

Rubric for Task Aa, Ba, and Ca

Steps with sub-steps of a proof by mathematical induction	Points
Basis step	1
Choosing the correct starting point, n_0	0.50
Correctly evaluating $P(n_0)$	0.50
Inductive step	7
Stating $P(n_0)$….. P(k)	1
Stating $P(k + 1)$	1
Substituting $P(n_0)$….. P(k) into $P(k + 1)$ to show $P(k) \rightarrow P(k + 1)$	1
Algebraic manipulations needed to show $P(k) \rightarrow P(k + 1)$	4
Writing the proof in the correct form	2
Invoking PMI	1
Stating the domain the introduced variable k	0.50
Explicitly writing that $P(k + 1)$ is to be shown assuming that $P(n_0)$….. P(k) is true	0.50

Total	10

Research Question 2

2) How do students' convictions differ for weak versus strong induction?

The extension of the WMRe framework from Gonzalez (2020) to analyze student convictions was used to answer Research Question 2. Results are recorded using **Table 3** and **Table 4**. This helped us determine the students' convictions for weak and strong induction for **Tasks Ab**, **Ac**, **Bb**, **Bc**, **Cb**, and **Cc**. Also, interactions for parts **b** and **c** for each of the convictions were also recorded for every task with a separate table.

Table 3

Student Convictions and Disbeliefs in Task 2a

	Conviction		Disbelief		
	Proof	Absolute	Unconvinced	Empirical	Total
Task Ab	-	N/A	-	-	-
Task Ac	-	-	N/A	-	-
Task Bb	-	N/A	-	-	-
Task Bc	-	-	N/A	-	-
Task Cb	-	N/A	-	-	-
Task Cc	-	-	N/A	-	-

Table 4

Student Interactions of Responses Between Task Ab and Task Ac

	Task Ab result			
Task Ac result	Proof conviction	Unconvinced disbelief	Empirical disbelief	Total
Proof conviction	-	-	-	-
Absolute conviction	-	-	-	-
Empirical disbelief	-	-	-	-
Total	-	-	-	-

Research Question 3

3) What factors or attributes do students have that relate to having difficulties with weak or strong induction? How does positive self-efficacy in using weak and strong induction significantly affect the students' performance?

To answer Research Question 3, the following tables were created to analyze student self-efficacy. **Table 5** is a frequency table created to analyze the occurrences of student self-efficacy for each statement. The data from **Table 5** were analyzed using the total score for each statement and comparing it to the performance scores from **Table 1** and **Table 2**. **Table 6** was created to analyze the frequency of self-efficacy by gender. The average female score and the average male score were analyzed to make comparisons.

Table 5

Frequency of Mathematical Induction Self-Efficacy by Statement

Score	Proof Statement 1	Strong Induction Statement 2	Weak Induction Statement 3	Correct Form Statement 4	Basis Statement 5	Inductive Statement 6
1	-	-	-	-	-	-
2	-	-	-	-	-	-
3	-	-	-	-	-	-
4	-	-	-	-	-	-
5	-	-	-	-	-	-

Table 6

Frequency of Mathematical Induction Self Efficacy by Sex

Score	Female	Male
1	-	-
2	-	-
3	-	-
4	-	-
5	-	-

Chapter Summary

In this chapter, methods of analysis were outlined for students' difficulties with weak and strong induction, students' convictions of proofs, and students' self-efficacy and the association with student performance. The E-a framework was used to score and analyze student difficulties for strong and weak induction. The WRMe framework in Gonzalez (2020) was used to measure student convictions for weak and strong induction. Finally, a Likert-scale was used to analyze self-efficacy of strong and weak mathematical induction. In the next chapter, the data analysis with the study results will be presented.

CHAPTER IV

DATA ANALYSIS

Research Question 1

Research Question 1 asked: "What difficulties do undergraduate mathematics students have doing proofs when using the method of weak induction versus strong induction?". Using Gonzalez (2020) rubric, students' answers were scored and the means for each task were calculated. The means for each task can be seen in **Table 7** for **Task Aa**, and **Table 8** for **Task Ba** and **Ca**.

Table 7

Mean scores for Task Aa

Steps with sub-steps of a proof by mathematical induction	Points
Basis step	0.80
Choosing the correct starting point, n_0	0.40
Correctly evaluating $P(n_0)$	0.40
Inductive step	0.80
Stating $P(k)$	0.40
Stating $P(k + 1)$	0.40
Substituting $P(k)$ into $P(k + 1)$ to show $P(k) \rightarrow P(k + 1)$	0
Algebraic manipulations needed to show $P(k) \rightarrow P(k + 1)$	0
Writing the proof in the correct form	1
Invoking PMI	0.40
Stating the domain the introduced variable k	0.40
Explicitly writing that $P(k + 1)$ is to be shown assuming that $P(k)$ is true	0.20

Total	2.3

Table 8

Mean scores for Task Ba and Ca

Steps with sub-steps of a proof by mathematical induction	Points	Points
Basis step	0.30	0.30
Choosing the correct starting point, n_0	0.10	0.20
Correctly evaluating $P(n_0)$	0.20	0.10
Inductive step	1	0.90
Stating $P(n_0)$….. P(k)	0.20	0.20
Stating $P(k + 1)$	0.20	0.20
Substituting $P(n_0)$….. P(k) into $P(k + 1)$ to show $P(k) \rightarrow P(k + 1)$	0.20	0.10
Algebraic manipulations needed to show $P(k) \rightarrow P(k + 1)$	0.40	0.40
Writing the proof in the correct form	0.70	0.70
Invoking PMI	0.40	0.40
Stating the domain the introduced variable k	0.20	0.20
Explicitly writing that $P(k + 1)$ is to be shown assuming that $P(n_0)$….. P(k) is true	0.10	0.10
Total	2	1.9

To answer the first research question about student difficulties with weak and strong induction we took a look at the mean scores for the components of the submitted proofs of the two types of induction in Table 7 (weak induction) and Table 8 (strong induction). We compared those scores with corresponding scores for the interviewed students in Table 9 (weak induction) and Table 10 (strong induction).

In **Table 7**, showing results from **Task Aa** (weak induction) the basis step the difficulties tended to be the same for the correct evaluation of the starting point and selecting the start point. The primary difficulty students had with weak induction, came from the inductive step. The main difficulties of the inductive step came from substituting $P(k)$ into $P(k + 1)$ to show $P(k) \rightarrow P(k + 1)$ and the algebraic manipulations needed to show $P(k) \rightarrow P(k + 1)$. Students seemed to have fewer difficulties with putting the proof in the correct form. Students seemed to be able to recognize the domain of validity.

The data in **Table 8** was analyzed to find the difficulties students have with strong induction when completing **Tasks Ba** and **Ca**. The inductive step once again was the primary difficulty that students were having for strong induction.

So, how did the weak induction results compare to the strong induction results? Students had a better average for the starting point of weak induction as opposed to the strong induction. Students had less difficulty stating the hypothesis for weak induction. Students had more difficulties with substituting and algebraic manipulation for strong induction. For making the proofs in the correct form students seemed to perform better on the weak induction. These observations support the fact that strong induction had more difficulties compared to weak induction.

Table 8 for strong induction and **Table 7** for weak induction show that, overall, students had the most difficulties with strong induction. This means students struggled more to complete a strong induction proof. Students also struggled the most with the inductive step for both weak and strong induction. This means that students may find the inductive step of the proof more difficult to complete than the other steps. Students found the basis step and the correct form easier to complete since they had fewer difficulties with these steps.

Students performed better on weak induction compared to strong induction overall and they found that the basis step was easier to complete for weak induction as well as the correct form. This means when it comes to choosing a start part students tend to do better for weak induction. They also tend to do better evaluating the starting point for weak induction. Students found invoking PMI equally as easy for both strong and weak induction. They found that stating the domain for k was easier for weak induction. They also found that explicitly writing that $P(k + 1)$ is to be shown assuming that $P(k)$ is true was easier than explicitly writing that $P(k + 1)$ is to be shown assuming that $P(1),...,P(k)$ are true.

Results from interview

The interview of the two students, S1 and S2, showed that the students had fewer difficulties when probed and asked to answer the tasks. Results can be seen in **Table 9** and **Table 10**.

Table 9

Mean Scores for Task Aa With Interview Discussion

Steps with sub-steps of a proof by mathematical induction	Points
Basis step	1
Choosing the correct starting point, n_0	0.50
Correctly evaluating $P(n_0)$	0.50
Inductive step	7
Stating $P(k)$	1
Stating $P(k + 1)$	1

Substituting $P(k)$ into $P(k + 1)$ to show $P(k) \rightarrow P(k + 1)$	1
Algebraic manipulations needed to show $P(k) \rightarrow P(k + 1)$	4
Writing the proof in the correct form	2
Invoking PMI	1
Stating the domain the introduced variable k	0.50
Explicitly writing that $P(k + 1)$ is to be shown assuming that $P(k)$ is true	0.50
Total	10

Table 10

Mean Scores for Task Ba and Ca With Interview Discussion

Steps with sub-steps of a proof by mathematical induction	Points	Points
Basis step	1	0
Choosing the correct starting point, n_0	0.50	0
Correctly evaluating $P(n_0)$	0.50	0
Inductive step	7	7
Stating $P(n_0)$….. P(k)	1	1
Stating $P(k + 1)$	1	1
Substituting $P(n_0)$….. P(k) into $P(k + 1)$ to show $P(k) \rightarrow P(k + 1)$	1	1
Algebraic manipulations needed to show $P(n_0)$… P(k) $\rightarrow P(k + 1)$	4	4
Writing the proof in the correct form	2	2
Invoking PMI	1	1
Stating the domain the introduced variable k	0.50	0.50

Explicitly writing that $P(k + 1)$ is to be shown assuming that $P(n_0)$….. P(k) is true	0.50	0.50
Total	10	9

They were able to get nearly full points for every category except the basis step where they did not receive any points for one of the strong induction tasks. Students during the interview were able to get the right starting point for **Task Ba**, but not **Task Ca**. They also did not complete the base step even with an incorrect starting point. There were some notable difficulties stated by the students themselves during the interview.

The students scored a lot higher because during the interview they were able to differentiate the tasks. The wording of the problems may have left students more confused without an interview to receive some type of probing for each task. The answers for the two students who were interviewed were accounted separately, but also as a mean of those two scores, since they were interviewed together, and they discussed about the tasks. Table 9 shows the results from the interview scores for **Task Aa**, and Table 10 shows the results from the interview scores for **Tasks Ba** and **Ca**.

Observer Notes from the Interview

Observer notes were written down as the students responded to each question. The questions probed about the **Tasks Aa**, **Ab**, **Ac**, **Ba**, **Bb**, **Bc**, **Ca**, **Cb**, and **Cc**, as seen in **APPENDIX A.** Results from the interview were significantly higher in performance when compared to the results when the students completed the tasks without interviews.

One of the students, S1, noted how it was "Easier to complete the proofs when we discuss and establish an understanding of what each task was asking." During the interview, S2 (the second student) said they initially wanted to do all the proofs by weak induction. When completing the tasks, they had a more algebraic approach to understanding the breaking of the chocolate bar. S1 tried to use diagrammatic pictures of the candy bars for each task to demonstrate the process of breaking the bars into individual pieces.

To start the interview, students were read out loud **Task Aa** and told to talk aloud while they worked through the proof. On this first task of weak induction, both S1 and S2 agreed on the correct starting point and clearly stated the basis step of the proof. When they were asked what to do next, both students knew to do the inductive step. S2, when asked about the inductive step, knew how to set up the assumption. S2 then started to do the algebraic manipulation out loud and got a little bit confused or lost in their words. So, the researcher wrote down what they were trying to explain mathematically. Their substitution of the pieces into the breaks was correct but some of the algebra was incorrect. Once seen on paper, S2 was able to correct their algebraic manipulation to prove that the task was true for k+1 with help from S1. S1 realized and stated that the initial break of the chocolate bar was not accounted for. They both proceeded to invoke PMI and finished the first proof. Both students were then asked what difficulties they had. S2 stated that they had a hard time remembering the initial break. S1 said they had difficulties when it came to understanding the problem.

The students were then read out loud **Task Ba** and then asked to do the proof for strong induction. S1 and S2 both agree that the basis step was the same. S2 then proceeded to state the assumption correctly for strong induction. Again, both students agreed that this was correct. They both knew and declared that the next step was showing it was true for k+1. They were able to use similar algebraic manipulations from the first task to show how the pieces can be substituted for the number of breaks. They used a little bit of backward thinking because they knew what their end goal was and manipulated the algebra to reflect their goal.

For the final strong induction **Task Ca**, students were read the problem once again. They were then asked to prove the statement using mathematical induction. They realized that the breaking was different but had the same number of breaks as before in the past two tasks. Neither of the students picked the correct starting point or correctly evaluate the starting point. They did use the correct assumption and they did the induction step correctly together. They also finished the proof by invoking PMI.

Both students were asked for difficulties for **Task Ca**. S1 said it was easier to prove **Task Ca** after proving **Ba**. S2 said that their difficulties were that they wanted to use weak induction.

Research Question 2

Research question 2 asked: "How do students' convictions differ for weak versus strong induction?" **Table 11** shows the frequency of conviction for each of the three tasks. Part b of each task had a number within the domain and part c had a number outside of the domain. The interactions between parts c and b of each of the three tasks were calculated again using a frequency table. **Table 12** represents the interaction between **Tasks Ab** and **Ac**. **Table 13** represents the interaction between **Tasks Bb** and **Bc**. **Table 14** represents the interaction between **Tasks Cb** and **Cc**.

Table 11

Student Convictions and Disbeliefs in Task Ab, Ac, Bb, Bc, Ab, and Ac

	Conviction		Disbelief		
	Proof	Absolute	Unconvinced	Empirical	Total
Task Ab	5	N/A	-	-	-

Task Ac	2	2	N/A	-	1
Task Bb	5	N/A	-	-	-
Task Bc	2	3	N/A	-	-
Task Cb	5	N/A	-	-	-
Task Cc	2	2	N/A	-	1

Table 12

Student Interactions of Responses Between Task Ab and Task Ac

	Task Ab result			
Task Ac result	Proof conviction	Unconvinced disbelief	Empirical disbelief	Total
Proof conviction	2(40%)	-	-	2(40%)
Absolute conviction	2(40%)	-	-	2(40%)
Empirical disbelief	1(20%)	-	-	1(20%)
Total	5(100%)	-	-	5

Table 13

Student Interactions of Responses Between Task Bb and Task Bc

	Task Bb result			
Task Bc result	Proof conviction	Unconvinced disbelief	Empirical disbelief	Total
Proof conviction	2(40%)	-	-	2(40%)
Absolute conviction	3(60%)	-	-	3(60%)

Empirical disbelief	-	-	-	
Total	5(100%)	-	-	5

Table 14

Student Interactions of Responses Between Task Cb and Task Cc

	Task Cb result			
Task Cc result	Proof conviction	Unconvinced disbelief	Empirical disbelief	Total
Proof conviction	2(40%)	-	-	2(40%)
Absolute conviction	2(40%)	-	-	2(40%)
Empirical disbelief	1(20%)	-	-	1(20%)
Total	5(100%)	-	-	5

Results for conviction show that all students have proof conviction for all three tasks for a statement $P(n)$, with a value n within the domain of validity. So, this means for strong and weak induction if a value n is given within domain of validity, students believe the statement is true and they believe in their proof.

Answers varied when the students were asked if the statement was true for a number outside of the domain of validity. For weak and strong induction **Tasks Ac** and **Cc** had the same results. A couple of students had proof conviction meaning that they understand the domain of validity and they believe in their proof but need to verify if it's true for the given value. A couple of students had absolute conviction meaning they think that the statement would be false for a number outside of the domain of validity and they believe in the values the proof showed. Only one student had empirical disbelief for **Tasks Ac** and **Cc**, meaning that they did not understand the statement and they incorrectly substituted a value into the statement that contradicted their own proof. **Task Bc** for strong induction had one difference, with no students having empirical disbelief and one more student for absolute conviction compared to **Task Ac** and **Cc**.

For interactions with convictions for weak induction, 40% of students had proof conviction for both **Tasks Ab** and **Ac**. This means that 40% of students were consistent with their proof conviction and they believe in their weak induction. So, they understood how to handle statements $P(n)$ with values n that are inside and outside of the domain of validity. Also, 40% of students had proof conviction for **Task Ab** and absolute conviction for **Task Ac**. This means that students believed in their proof for values n within the domain of validity and they knew the statement was true for those values. This also means that if the value n was outside of the domain of validity, they believed the statement was false because they believed in their proof for values in the given domain. 20% of students had proof conviction for **Task Ab** and empirical disbelief for **Task Ac**. This means they thought that the statement was true for values of n within the domain of validity, and they believed in their proof. But, for a value n outside of the domain of validity, they did not understand and substituted incorrectly and ended up contradicting their proof. The same results can be seen for **Task Cb** and **Cc** interactions for strong induction.

For one of the strong inductions tasks the interactions were slightly different. 40% of students had proof conviction for **Task Bb** and proof conviction for **Task Bc**, meaning that those students believed in their own proof and found the statement $P(n)$ to be true for values of n inside and outside of the domain of validity. 60% of students had proof conviction for **Task Bb** and absolute conviction for **Task Bc**. This means that they believed in their proof and believed the statement $P(n)$ was true for values of n within the domain. For values of n outside of the domain, students thought the statement was false because they believed in their proof and thought verification was not needed for values *n* outside of the domain of validity. Student convictions for the interview were recorded and are included in the results with the convictions.

Research Question 3 Data Analysis

Research question 3 asked "What factors or attributes do students have that relate to having difficulties with weak or strong induction? How does positive self-efficacy in using weak and strong induction significantly affect the student's performance?"

Students were given a score on a scale of 0 to 5 with 0 for "strongly disagree "and 5 for "strongly agree." Statements of self-efficacy were listed in Chapter III. The frequency for the score of all students' responses was recorded for each of the six self-efficacy statements on **Table 15** below.

Table 15

Frequency of Mathematical Induction Self-Efficacy by Statement

Score	Proof Statement 1	Strong Induction Statement 2	Weak Induction Statement 3	Correct Form Statement 4	Basis Statement 5	Inductive Statement 6
1	-		-	1		-
2	1	2	1	-	-	2
3	-	2	-	1	-	-
4	2	1	2	3	-	3
5	2	-	2	-	5	-
	20	14	20	16	25	16

We can see from **Table 15** that the self-efficacy score was the lowest for proving strong induction. The second lowest self-efficacy score was for doing mathematical induction in the correct form and the inductive step. The self-efficacy score was the highest for proving the basis step. The second highest self-efficacy score was for completing a mathematical proof in general and completing a weak induction proof.

Overall, when referring back to performance, if we look at student scores for the mathematical induction tasks student got 50% of the possible points available for the basis step for weak induction(**Task Aa**). Students earned 30% of the possible points for the basis step for both strong induction problems (**Task Ba** and **Ca**). The percentage of points earned for the inductive step for weak induction (**Task Aa**) is 11%, for strong induction **Task Bb** is 14%, and for **Task Cc** is 13%. For the correct form for weak induction (**Task Aa**), the percentage of points earned is 50%. For the correct form for strong induction (**Task Ba**), the percentage of points earned is 35%, and for **Task Ca** is 35% as well. Although, the difficulties seem to be more for weak induction in the inductive step, the scores were very close together and fairly low for both forms of induction. But, overall, most difficulties were with strong induction since the percentage of points was significantly lower for the basis step as well as the form.

Performance Percentage Score Comparison to Self-Efficacy Score

The lowest self-efficacy score of all students was 11 for the ability to prove strong induction. Also, students had lower performance on the strong induction tasks. Overall, strong induction had a lower performance score compared to weak induction performance. Students had 20% percent less points for the basis step and 15% less points for the correct form. The score for strong induction was about 4% higher for the inductive step, compared with the weak induction, but the difference was not significant.

The highest self-efficacy score went to the ability to complete the basis step, with a total score of 25 when all the students' points were added together. Students performed better with the basis step than with the inductive step. They had a score of 50% overall score for weak induction and a score of 30% overall for both strong induction tasks.

Students got a total self-efficacy score of 19 for weak induction. Their total percentages for possible points were higher, as mentioned before, for weak induction.

The self-efficacy score total for the correct form was 16. The performance percentage of possible points that can be earned for the correct form was 50 percent for **Task Aa**. The percentage was 35% for the correct form for both **Tasks Ba** and **Ca**.

Analysis of the Confidence Statements

The data for the individual confidence statements are ordinal, so we can pairwise test the statements for a greater degree of confidence using the Mantel-Haenszel Test for Linear Trend (STAT 504, 2022). Table 17 shows the correlation coefficient and p-value that there is no linear trend from one statement to the other.

Table 17

Analysis of Confidence Statements Using Mantel-Haenszel Test for Linear Trend

Score	Proof Statement 1	Strong Statement 2	Weak Induction Statement 3	Correct Form Statement 4	Basis Statement 5	Inductive Statement 6
1	-	-0.53,0.105	0,1	-0.33,0.31	-0.54,0.103	-0.35,0.28
2	-	-	0.53,0.105	0.2,0.54	0.90,0.0068	0.22,0.50
3	-	-	-	-0.33,0.31	-0.54,0.103	-0.35,0.28
4	-	-	-	-	0.74,0.027	0,1
5	-	-	-	-	-	-0.79,0.017

A p-value less than 0.05 means that there is a statistically significant comparison between the two statements of confidence. A p-value of approximately 0.10 shows weak statistical significance. A negative correlation coefficient means that there is greater confidence in the first statement.

Comparisons with $p < 0.05$ there is greater confidence in Statement 5 (prove a basis step) than in Statements 2 (proofs by strong induction), 4 (inductive proofs in the correct form), and 6 (prove an inductive step). Comparisons with p-value of approximately 0.10 show weak agreement with Statements 1 (complete a mathematical proof) and 3 (proof by weak induction) showing greater confidence that Statement 2. Such comparisons also show weak agreement that Statement 5 shows greater confidence than Statements 1 and 3.

This shows that the population of mathematics majors has the greatest confidence in being able to complete the basis step of a proof by mathematical induction. There is less confidence in being able to complete mathematical proofs in general and proofs by weak mathematical induction. Finally, mathematics majors have the least confidence in completing proofs by strong induction, completing the inductive step, and ensuring that the proofs are in the correct form.

Table 16 was used to measure the frequency of scores for each gender to make comparisons.

Table 16

Frequency of Mathematical Induction Self-Efficacy by Gender

Score	Female	Male
1	1	-
2	5	1
3	-	3

4	4	7
5	2	7
	47	74

The data in **Table 16**, female versus male frequency for self-efficacy score for each statement, were analyzed to see if there is a significant difference in self-efficacy score for males versus females. The average female score and the average male score were analyzed to make comparisons. Females got an average of 23.5 overall self-efficacy score while males got 24.66 as their average. This was out of a possible 30 points. When it comes to the performance of all mathematical induction proofs for Tasks Aa, Ba, and Ca females scored an average of 5.75 out of 10 possible points and males scored an average of 6.17 out of 10 possible points.

The data in **Table 16**, female versus male frequency for self-efficacy score for each statement, were analyzed to see if there is a significant difference in self-efficacy score for males versus females using a chi-squared test. The p-value was 0.3977, so there is no significant difference. But, it might show possible signs for future research in larger sample numbers.

Chapter Summary

In this chapter, data analysis for student difficulties with weak and strong induction, students' convictions of proofs, and students' self-efficacy and the association with student performance was discussed. Comparisons were looked at in the results of strong and weak mathematical induction performance. Comparisons were looked at for student convictions with strong and weak induction as well. Finally, comparisons to self-efficacy scores and student performance were made with strong and weak induction. In the next chapter, the discussion and conclusions of the study results will be presented.

CHAPTER V

DISCUSSION AND CONCLUSIONS

Introduction

This chapter includes a discussion of the results and how they address each research question.

Research Question 1 Discussion

Research Question 1 compares student performance using weak verse strong induction. "What difficulties do undergraduate mathematics students have doing proofs when using the method of weak induction versus strong induction?"

As we saw in the data analysis, students had more difficulties with the inductive step overall for mathematical induction. Gonzalez (2020) was able to study student difficulties with weak induction and he found that students had the least difficulties with the basis step which was similar to this study's findings. The new finding was that the students had more difficulties with the basis step for strong induction compared to weak induction. Students struggled more with putting proofs in the correct form for weak induction in Gonzalez's study (2020). They had the most difficulties with the inductive step. For our study students struggled the most with the inductive step for both weak and strong induction compared to the correct form. For the correct form students had more difficulties with strong induction as opposed to weak induction. These differences might be because the sample size was smaller than Gonzalez (2020) study. It would be interesting to make strong and weak comparisons with a larger group of participants in future studies.

New findings from this study also showed that students overall had more difficulties with strong induction compared to weak induction. This comparison hasn't been analyzed before in any other study. So, it would be interesting to continue to compare these difficulties with larger sample sizes when possible.

Research Question 2 Discussion

The second research question addressed proof of conviction for the two forms of induction. "How do students' convictions differ for weak versus strong induction?"

Gonzalez (2020) showed we were able to see what convictions students had for weak induction and the interactions of their convictions. Our study compared student convictions with weak and strong induction, which hasn't been seen before. Results also showed that overall students will have the same convictions for weak induction that they have for strong induction. The differences occurred when they are presented with a value within the domain of ~~the variable~~ validity and a value outside of the domain of the variable. Every student had proof conviction for all tasks for a value within the domain of validity. This could have differed from Gonzalez's study for many reasons. Possibly because the students were mathematics majors. Maybe it could have also been from his larger sample size. Similarly, in Gonzalez's (2020) study for the value outside of domain of validity, students had primarily a mix between proof and absolute conviction.

Our study showed overall that student convictions would be the same for weak and strong mathematical induction. This has not been analyzed before and would be interesting to continue researching and performing studies. If students have the same conviction for strong and weak induction that means they are consistent with their beliefs and the type of proof by induction does not affect their convictions.

Research Question 3 Discussion

The third research question considered an external factor for student performance in using the two forms of induction. "What factors or attributes do students have that relate to having difficulties with weak or strong induction? How does positive self-efficacy in using weak and strong induction significantly affect the student's performance?"

Self-efficacy appears to be linked to performance as we saw in the literature review. It also has been linked to mathematical performance. Hackett and Betz (1989) had results showing higher self-efficacy called for higher performance specific to mathematics skills. But, it hasn't been looked at for specifically mathematical induction. Self-efficacy also hasn't been compared to performance for strong verse weak induction. The study in this thesis showed how weak induction had a higher self-efficacy score and a higher performance while strong induction had a lower self-efficacy score and lower performance. Hackett and Betz (1989) also found that males tend to have higher self-efficacy in STEM-related skills, compared to women. This study found weak evidence for mathematical induction since women scored themselves lower overall for self-efficacy while males scored themselves higher overall for self-efficacy. This is related to lower performance from females overall compared to males for mathematical induction. Performance, according to Hackett and Betz (1989), was lower for females as well because of their self-efficacy, which in turn, may be linked to social anxieties and other factors.

Implications for Teaching

This study is important to look at because, in education, strong mathematical induction is not typically analyzed for difficulties at the college level. Gonzalez (2020) was able to study undergraduate students and their difficulties with weak induction. This study, on the contrary, focused on mathematics majors and their difficulties instead of a more generalized population.

This study will add to the few studies for mathematical induction, conducted with mathematics majors. Most studies focus on weak induction. Strong induction has not been compared or analyzed for difficulties.

To make improvements to the teaching of mathematical induction, emphasis on the inductive step would help possibly avoid difficulties that students are having. If we focus on the main area of concern and start there, we can avoid difficulties and overall improve the teaching of mathematical induction.

The study included the reflection of two students who promoted each other's understanding of the tasks and how to complete them. To improve the teaching of mathematical induction there could be a discussion among peers. As we saw in the interview, performance was significantly better than when students had to perform the tasks on their own with no discussion involved. Students in the interview were able to talk through each of the steps and build confidence in their work.

This study suggests a link between self-efficacy and student performance for undergraduate mathematics students for mathematical induction. Teachers may use self-efficacy to get an idea of areas of improvement the student might have. If students have a lower self-efficacy, as they did with this study for strong induction, teachers can focus on that content first to target the difficulties and overall help students improve their performance even at the undergraduate level.

When it comes to conviction, if the teacher notices that students have proof conviction for a value within the domain of validity, then they might want to emphasize and analyze convictions when students are given a value outside of the domain. So, maybe clarification is needed to help students understand and form conviction from their new knowledge.

Limitations of the study

One limitation of the study was getting participants that not only met the criteria of the study but wanted to participate. The target population, math majors, is relatively small, to begin with. Interaction with possible participants was limited because of COVID. Only 5 participants in total completed the instruments.

Another possible limitation is the fact that the inductive tasks were chosen with the least amount of rigorous algebraic manipulations so that the study could focus primarily on mathematical induction itself and the weak and strong inductive comparison. So, comparisons to studies with more algebraic manipulation may be limited.

Future Research

While this study did give indications of difficulties students had with strong and weak induction, future studies with better recruitment could use standardized research designs, or teaching experiments such as in Norton, Arnold, and Koshushkin (2022). The tasks could be expanded to include those with algebraic components.

Since this study suggests there is a connection between gender and self-efficacy for mathematical induction, further research could confirm this connection. This could contribute to an understanding of the relationship between gender and mathematical self-efficacy for other topics in upper-level mathematics.

REFERENCES

Avital, S., & Hansen, R. (1976). Mathematical Induction in the Classroom. *Educational Studies in Mathematics*, *7*(4), 399-411. Retrieved September 8, 2021, from http://www.jstor.org/stable/3481946

Avital, S., & Libeskind, S. (1978). Mathematical Induction in the Classroom: Didactical and Mathematical Issues. *Educational Studies in Mathematics*, *9*(4), 429-438. Retrieved September 1, 2021, from http://www.jstor.org/stable/3481911

Baker, A. (2010). Mathematical induction and explanation. *Analysis*, *70*(4), 681-689. Retrieved January 23, 2021, from http://www.jstor.org/stable/40930280

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Benson, J. (1989). Structural components of statistical test anxiety in adults. *Journal of Experimental Education*, *57*, 247-261.

Bogomolny, A., & Taleb, N. N. (2020).*Cut the knot: Probability riddles*. Wolfram Media.

Brown, I., Jr., and Inouye, D. K. (1978). Learned helplessness through modeling: The role of perceived similarity in competence. *Journal of Personal and Social Psychology*, . *36*, 900-908.

Ernest, P. (1984). Mathematical Induction: A Pedagogical Discussion. *Educational Studies in Mathematics*, *15*(2), 173-189. Retrieved January 23, 2021, from http://www.jstor.org/stable/3482247

Goldin, A. (2000). A scientific perspective on structured, task-based interviews in mathematics education research. In Kelly A. E. & R. A. Lesh (Eds.) *Handbook of research design in* *mathematics and science education* (pp. 517-545). Mahwah, New Jersey: Laurence Erlbaum Associates.

Gonzalez, M. (2020). Undergraduate students' difficulties and convictions with mathematical induction. Unpublished thesis. Texas A&M University-Corpus Christi. Retrieved August 20, 2021 from https://tamucc-ir.tdl.org/handle/1969.6/89226

Hackett, G., & Betz, N. (1989). An Exploration of the Mathematics Self-Efficacy/Mathematics Performance Correspondence. *Journal for Research in Mathematics Education*, *20*(3), 261-273. doi:10.2307/749515

Harel, G. (2001). The development of mathematical induction as a proof scheme: A model for DNR-based instruction. In S. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory* (pp. 185-212). New Jersey: Ablex Publishing Company.

Harel, G. & Sowder, L. (2007) Towards a comprehensive perspective on the learning and teaching of proof. In Lester, F. (Ed.) *Second Handbook of Research on Mathematical Teaching and Learning* (pp. 805-842). Washington, DC: NCTM.

Lent, R. W., Brown, S. D., & Larkin, K. C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. Journal of Counseling Psychology, *33*, 256-269.

Liu, Xing & Koirala, H. (2009). The Effect of Mathematics Self-Efficacy on Mathematics Achievement of High School Students. NERA Conference Proceedings 2009.

Norton, A., Arnold, R., Koskushkin, V. & Tiraphatna, M (2022). Addressing the Cognitive Gap in Mathematical Induction. *International Journal of Research in Undergraduate Mathematics Education*, https://doi.org/10.1007/s40753-022-00163-2

Polya, G. (1954). *Mathematics and plausible reasoning. I. Induction and analogy in mathematics. II. Patterns of plausible inference.* Princeton University Press.

Schach, A. (1958). Two Forms of Mathematical Induction. *Mathematics Magazine*, *32* (2), 83-85. doi:10.2307/3029535

Schunk, D. H. (1989). Self-efficacy and achievement behaviors. Educational Psychology Review, *1*(3), 173–208. https://doi.org/10.1007/BF01320134

Siegel, R. G., Galassi, J. P., & Ware, W. B. (1985). A comparison of two models for predicting mathematics performance: Social learning versus math aptitude-anxiety. *Journal of Counseling Psychology*, *32*, 531-538

STAT 504. (2022). Department of Statistics, Penn State University. Retrieved November 12, 2022, from https://online.stat.psu.edu/stat504/lesson/4/4.4

Stylianides, G., Stylianides, A., & Philippou, G. (2007). Preservice teachers' knowledge of proof by mathematical induction*. Journal of Mathematics Teacher Education*, *10*(3), 145–166.

The National Counsel of Teacher of Mathematics. (n.d.). Nctm.org. Retrieved November 22, 2022, from https://www.nctm.org/Standards-and-Positions/Principles-and-Standards/Process/

Temple, G. (1957). The Growth of Mathematics: Presidential Address to the Mathematical Association, January 1957. The Mathematical Gazette, 41(337), 161–168. https://doi.org/10.2307/3609185

Usher, E. L., & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. *Contemporary Educational Psychology*, *34*(1), 89101, https://doi.org/10.1016/j.cedpsych.2008.09.002

Weber, K., & Mejia-Ramos S, J. P. (2015). On Relative and Absolute Conviction in Mathematics. *For the Learning of Mathematics*, *35*(2), 15–21. http://www.jstor.org/stable/44382751

Zimmerman, B. J., & Ringle, J. (1981). Effects of model persistence and statements of confidence on children's self-efficacy and problem solving. *Journal of Educational Psychology*,*73*,485-493.

APPENDIX A

INSTRUMENTS

Directions: Please take a moment to complete section 0, 1 and 2. There are short response and circle choice questions. You will then be asked to do an oral interview with questions about mathematical induction based on previously given tasks. After the oral interview you will then go to section 3 and if you would like to change any responses from section 2 you may do so there.

Section 0:

Age:

Sex:

Year of Undergraduate:

In how many college courses have you learned about or used mathematical Induction?

Were you aware of mathematical induction before college?

Section 1:

1) What is a mathematical proof?

2) What is a proof by mathematical induction? What are the steps?

3) What is a proof by strong induction?

4) Is there a significant difference between strong and weak induction?

Section 2:

Indicate your agreement with the following statements about typical work you have done for classes by circling an answer for each statement.

I am confident that I can complete mathematical proof.

Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree

I am confident that I can complete mathematical proof by strong induction.

Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree

I am confident that I can complete a mathematical proof by weak induction.

Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree

I am confident in my ability to do a proof by mathematical induction in the correct form.

Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree

I am confident in my ability to prove the basis of induction.

Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree

I am confident in my ability to prove the induction step.

Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree

Before the oral interview please briefly look over these definitions and concepts below.

Mathematical induction is a specific method of proof for statements that are quantified for the set of positive integers or specific subsets of that set.

The **principle of weak mathematical induction** can be described as the following; If 1 has property P, and if any n having property P implies that n+1 has property P, then every n has property P.

Similarly, **the principle of strong mathematical induction** states that if 1 has property P, and if for all integers from 1 to n, having property P implies that n+1 has property P, then every n has property P.

The **difference between weak induction and strong induction** only appears in induction hypothesis. In weak induction, we only assume that a particular statement holds at k-th step, while in strong induction, we assume that the particular statement holds at all the steps from the base case to k-th step

Tasks

Task A. You are given a candy bar with n pieces in one row. If $n > 1$, you break off pieces one at a time. How many breaks do you need to separate the original bar into n individual pieces?

A(n): Breaking a chocolate bar with $n \geq 6$ pieces into individual pieces requires n-1 breaks.

a) Prove the statement A(n) by mathematical induction for $n \in \mathbb{N}$, $n \geq 6$.

b) Is the statement true for $n = 10$? Justify your answer.

c) Is the statement true for $n = 5$? Justify your answer.

Task B. You are given a candy bar with n pieces in one row. If $n > 1$, you break it into two separate, smaller bars, and then continue breaking the smaller bars into smaller bars or individual pieces until there are no bars with more than one piece. How many breaks in all are needed to separate the original bar into n individual pieces?

B(n): Breaking a chocolate bar with $n \geq 6$ pieces into individual pieces as described in this task requires n-1 breaks.

a) Prove the statement B(n) by mathematical induction for $n \in \mathbb{N}$, $n \geq 6$.
b) Is the statement true for $n = 10$? Justify your answer.
c) Is the statement true for $n = 5$? Justify your answer.

Task C. You are given a candy bar with n pieces in one row. If $n = 2$, break the bar into two individual pieces. If $n > 2$ you break the candy bar into two parts, one with 2 pieces and one with the remaining n-2 pieces. Then you continue breaking off 2-piece parts and breaking the 2-piece parts into 2 separate individual pieces. How many breaks (as close to even as possible) in all are needed to separate the original bar into n individual pieces?

C(n): Breaking a chocolate bar with $n \geq 6$ pieces into individual pieces as described in this task requires n-1 breaks.

a) Prove the statement C(n) by mathematical induction for $n \in \mathbb{N}$, $n \geq 6$,

b) Is the statement true for $n = 10$? Justify your answer.

c) Is the statement true for $n = 5$? Justify your answer.

Oral Section of Interview

1. What is your meaning of [name a specific mathematical concept?

2. What technique of proof did you use for each task (Task Aa, Ba, and Ca)? Can you justify your choice?

   Strong Mathematical Induction or Weak Mathematical Induction.

3. Can you identify the hypothesis/conclusion?

4. What happens with your proof construction if we remove this part of the hypothesis…, does your proof still hold true?

5. Where exactly in your proof did you use this part of the hypothesis…?

6. Can you justify how you advanced in your proof from this step… to this step…?

7. You use this statement … in your proof, can you justify it? Why does it hold?

8. What was difficult about each proof (Task Aa, Ba, and Ca)? Why?

9. Would the statement still be true for this new domain? Why?

10. Could you identify a new domain for this statement that is still true? Why does it still work?

11. Is the statement true for [this specific/special case]?

12. Imagine you had a friend who didn't understand [this part of the proof], how would you explain it to him or her?

13. How could you generalize/specialize this proof?

14. (Before proving the statement) Do you believe the statement is true?

15. How could you prove this statement a different way?