PREDICTING STUDENTS' APPROACH TO KNOWLEDGE BY ANALYZING MISCONCEPTIONS AND ERRORS STUDENTS MAKE WRITING PROOFS

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF SCIENCE

IN THE GRADUATE SCHOOL OF THE

TEXAS WOMAN'S UNIVERSITY

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

COLLEGE OF ARTS AND SCIENCES

BY

JENNIFER RICE, B.S.

DENTON, TEXAS

DECEMBER 2016

DEDICATION

For Nick. You have been with me every step of this journey. You never gave up on me, even when I was ready to give up on myself. You are my inspiration, my guide, my partner, and my love. With you by my side, I have the strength to accomplish anything.

ACKNOWLEDGEMENTS

I would like to thank Dr. Wheeler for giving me the opportunity to participate in this research. You knowledge, guidance, support, and friendship gave me the confidence to pursue this project. You have been an inspiration to me.

I would also like to thank Dr. Edwards for his unwavering patience and support. Without your professionalism and guidance, I could not have completed this journey. Dr. Marshall, your expertise has been an invaluable asset to this project. Thank you.

Finally, I would like to thank my darling children. You have all been supportive and loving, and I cannot wait to help you as you being your own journey. Everything I do is for you. I love you.

ABSTRACT

JENNIFER RICE

PREDICTING STUDENTS' APPROACH TO KNOWLEDGE BY ANALYZING MISCONCEPTIONS AND ERRORS STUDENTS MAKE WRITING PROOFS

DECEMBER 2016

For this quantitative research study, an existing proof rubric (Selden & Selden, 1987; Wheeler & Champion, 2011) for students' errors and misconceptions was used to examine students' proofs in Discrete Mathematics class. Data were collected from two classes spanning two semesters, and participants included 27 computer science majors and 35 mathematics majors. Through coding of student work, one misconception and two errors were identified and used to predict the students' thinking methodology and its impact on proof writing abilities.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Purpose of the Study

One of the most important skills taught in today's classroom is logic and reasoning skills (Gokkurt, Soylu, & Sahin, 2014; National Council of Teachers of Mathematics [NCTM], 2000; Weber, 2008). In mathematics, those skills are manifested in proof.

Some experts argue proof is a formal process, while others argue it is the product of creativity. Others still claim proof is a combination of both. Despite decades of research, there is no consensus in the mathematics-education community about what constitutes a mathematical proof (Hanna, 2000; Weber, 2008), but mathematicians accept the general characterization that proof begins with a set of axioms and follows well-defined principles and logical rules to a conclusion (Griffiths, 2000; Weber, 2008).

In the classroom, proof is the key to mathematical understanding (Hanna, 2000), and the measure of a student's mathematical understanding can be found within each student's construction of proof. Middle school, high school, and entry-level college mathematics classes tend to lack formalism within mathematical definitions and proofs (Gries & Schneider, 1995; Uhlig, 2002). When students reach college-level introductory proof-writing courses, they have little to no experience with proofs and mathematics rigor (Uhlig, 2002). Perhaps, this realization explains why many students think of proof as daunting and meaningless when it comes to establishing mathematical knowledge (de Villiers, 1991; Gokkurt, *et al*., 2014).

The National Council of Teachers of Mathematics, however, identifies proof as one of the top five process standards in the *Principles and Standards for School Mathematics* (NCTM, 2000). Proof is, in fact, a fundamental component of mathematical concept development, critical to both establishing the validity of an argument and passing on mathematical knowledge (Gokkurt, *et al*., 2002; Rav, 1999; Stylianides, 2007). As a result, proof instruction has been scrutinized in recent years, and attempts to understand and improve students' perceptions of proof and proof-writing abilities has been the subject of significant research. Through these investigations, the mathematics community has a deeper understanding of foundational deficiencies at all levels of education, as well as problems with reading, constructing, and validating proof.

Several studies identify specific errors and underlying misconceptions students make when attempting to construct proof (Gokkurt, *et al*., 2014; Selden & Selden, 1987, 2003; Wheeler & Champion, 2011). Within these studies, it is interesting to note that students who previously performed well in the algorithm-intensive courses that make up lower-division mathematics curriculum in most college-mathematics programs often realize diminished success when faced with the abstract logic and nontechnical complexity of proof writing (Selden & Selden, 1987; Wheeler & Champion, 2011).

Lower-division college mathematics curriculum is a mixture of mathematical-concept development and algorithm-rich problem-solving techniques. For the student, the challenge to reason through a mathematical idea or concept has not been issued. Instead, the task is to select the appropriate algorithm and apply the algorithm correctly. This type of coursework, while necessary, fails to develop reasoning skills or inspire the mathematical thinking required for writing proof (Selden & Selden, 1987).

Computer science students face a similar problem. The development of reasoning skills is emphasized in the computer science curriculum, yet algorithm-intensive processes embody the world of computer science and contribute largely to algorithmic thinking practiced by computer scientists (Knuth, 1974, 1985).

In 1985, mathematician and computer scientist, Donald Knuth, wondered if mathematicians and computer scientists apply different thinking processes to work within their respective fields. He conducted a study to compare algorithmic thinking and mathematical thinking and ultimately reported multiple differences in the way information is processed.

An algorithmic approach to information, while well organized, may lack complexity and vigor since algorithmic processes are fundamentally non-uniform. In contrast, mathematical thinking is often inefficient within complex and creative processes (Knuth, 1985). In his 1985 article, *Algorithmic Thinking and Mathematical Thinking*, Knuth refers to discipline-specific personality profiles while comparing algorithmic thinking with classical mathematical thinking as applied to formula manipulation, behavior of function values, abstract reasoning, generalization, and infinite dimensional spaces. His hypothesis was that individuals in various disciplines have different thought trajectories that chart their work. In other words, mathematicians view things differently than lawyers or culinary artists or writers. Similarly, experts in different fields of science have different mental processes and a different approach to knowledge.

Knuth's (1985) research established that individuals in the fields of computer science and mathematics each employ a variety of modes of thought, some of which overlap and some of which are unique to their respective disciplines. Computer scientists tend to be fundamentally adept to deal with process states, a skill intimately related to algorithmic thinking. Moreover, Knuth suggests that computer scientists are more inclined to manage a multitude of non-homogenous cases than a traditional mathematician, which further contributes to the notion of algorithmic thinking.

While computer scientists tend to lack determination with regard to uniformity, this can be explained by the fact they are frequently required to manage non-uniform concepts fluently. Using this model, distinctive personality profiles can be used to differentiate various areas of scientific thinking (Knuth, 1985).

The purpose of this study is to build on the idea of discipline-specific personality profiles as it relates to proof construction. In this study, proofs written by algorithm-minded students (computer science majors) and mathematically minded students (mathematics majors), each with algorithm-intensive and mathematical proof-writing course requirements, were compared to determine if a relationship exists between student major and the errors and misconceptions revealed when constructing proof in a junior-level Discrete-Mathematics class.

CHAPTER II

REVIEW OF LITERATURE

Proof has taken center stage in the mathematics community in recent years as a fundamental aspect in mathematics and a core concept in comprehensive and coherent mathematics curriculum (Gokkurt, *et al.*, 2014; Inglis & Alcock, 2012; NCTM, 2000; Selden & Selden, 2003; Weber, 2008). Leading experts in mathematics have studied, defined and redefined, and reimagined the role of proof in mathematics and mathematics education.

Accordingly, substantial research has been devoted to the pedagogical development of proof in the classroom (NCTM, 2000; Stylianides, 2007), as well as student comprehension of and ability to read and construct proof (Inglis & Alcock, 2012; Weber, 2008).

Previous studies range from surveying the degree of teacher engagement to examining student interpretation of the task of proof to observing validations of proof by seasoned mathematicians. The results highlight gaps and discrepancies in mathematical foundations and the knowledge of proof, as well as offer insight into possible pedagogical improvements.

In 2003, Selden and Selden examined the way students both generate proofs and reflect on the texts that establish truth and correctness of the proof. The exploratory study involved 80 university students majoring in mathematics and secondary mathematics education. The students were each taking a proof-writing course required for degree completion. During the study, students were asked to construct a proof based on a single theorem and then reflect on the proof as texts to determine correctness.

While constructing a proof is difficult in its own right, Selden and Selden (2003) illustrate that validating proof as texts goes beyond the rigor of constructing a proof. Validations redirect the burden of construction of meaning to the reader. Validations are believed to expand the reader's understanding of a theorem and tend to involve passionate consideration of both old and new ideas. It is that intense reflection that ultimately results in knowledge construction, thereby bridging the links between one's ideas.

Upon completion of individual interviews and data analysis, the researchers (Selden & Selden, 2003) concluded that students tended to focus on surface features of proofs opposed to the underlying logical structure and therefore cannot reliably execute a step-by-step validation of a proof without instruction. With minimal guidance, students who reflect upon each argument as part of validation demonstrated a significantly higher success rate.

Inglis and Alcock (2012) built on the Selden and Selden (2003) study, comparing the proof-validation behavior of beginning undergraduate students and research-active mathematicians. In the article, proof was referred to as an onerous process requiring an intimate knowledge of mathematical definitions and theorems, and the ability to reason through that knowledge. Each statement requires thorough evaluation until its validity can be established.

Previous findings from proof-validation studies have been based on written reports of practicing mathematicians (Weber & Mejia-Ramos, 2013) or verbal procedures which required students and mathematicians to speak their thoughts as part of the proof-validating process (Inglis & Alcock, 2012; Selden & Selden, 2003; Weber, 2008). For the Inglis and Alcock (2012) study, the authors recorded the eye movements of participants as they validated unsubstantiated proofs in addition to studying the results of the validations. Inglis and Alcock note that a proof-reader's behavior varies according to his or her objective. In other words, reading for validation, where the truth of a statement is under intense scrutiny, is quite different that reading for comprehension, where the truth of a statement is assumed true.

The Inglis and Alcock (2012) study participants included 18 first-year undergraduate students who had successfully completed two semesters of proof-based calculus and linear algebra, and 12 academic mathematicians from a high-ranked research-intensive university. The proofs used in the study were identical to the Selden and Selden (2003) work.

After analyzing eye-movement data, Inglis and Alcock (2012) found no significant difference in the average real-time distribution during validation between the two groups. The validations themselves, however, allowed the authors to substantiate Selden and Selden's (2003) claims that undergraduate students cannot consistently differentiate between invalid and valid arguments. Furthermore, the undergraduate students focused a disproportionate amount of time on formulas, suggesting the Selden and Selden surface-features claim is true. In contrast, the mathematicians' results revealed that the academics failed to agree on the validity of arguments, although it was unclear if the disagreements were mathematical in nature or merely style issues. Finally, the mathematicians put considerable effort into inferring embedded merits, whereas the undergraduate students did not.

Inglis and Alcock (2012) concluded their article highlighting the inherent difficulties that arise when students are unable to consistently and accurately read proof. Finally, they acknowledged a deficiency in mathematics education and highlighted a need to rethink pedagogical strategies and efforts.

In 2011, Wheeler and Champion deepened the understanding of students' errors and misconceptions when attempting to write proofs. In their mixed methods study, the authors examined one-to-one and onto proofs of 23 undergraduate students all taking an abstract algebra class. Errors and misconceptions were identified and coded using an adapted version of Selden and Selden's (1987) rubric of undergraduate students' proof misconceptions and errors.

The Wheeler and Champion (2011) study findings included that failing to define one or more variables within the proof, understanding how to utilize given information to write a correct mathematical proof, and distinguishing among different proof formats were the most common errors. Furthermore, when students were asked to apply the one-to-one and onto properties proved to homomorphism, the participants demonstrated considerable difficulty which suggests they memorized the procedures for writing specific types of proofs.

A 2014 qualitative proof-research study by Gokkurt, Soylu, and Sahin concentrated on university students studying physics, chemistry, and other fields of science. The researchers examined the proof-writing methods of 50 first-year college students studying science teaching. The randomly selected participants were challenged with three direct, one induction, and two geometry proofs. Student work was systematically analyzed using a rule-based coding rubric. Finally, the researchers compared and consolidated their results.

At the end of the year-long study, Gokkurt, Soylu, and Sahin (2014) concluded that while students employ multiple different strategies when attempting proof, the majority of them attempt to prove by example. In other words, the majority of students with a science background believe demonstrating the truth of a statement by inserting numerical values is adequate.

Another study (Stylianides, 2007) evolved the focus of proof development from student learning to teacher capability, with the author advocating a robust knowledge of proof is vital to proof cultivation among students. The Stylianides analysis emphasized a relationship between proof-and-proving and doing-and-knowing mathematics. Proofs are the foundation of mathematics. Proof is a fundamental element of developing, substantiating and conveying mathematical knowledge (Kitcher, 1984; Stylanides, 2007).

Before beginning the research, Stylianides (2007) acknowledged a critical shortfall in the universe of proof and called for supplementary research dedicated to abstracting a consistent and universal meaning of proof for all grade levels. For this project, he established a comprehensive multi-page definition of proof to prevent empirical arguments from being considered as proofs. Moreover, the author used the definition as a tool to conceptualize how its application could be used to support analysis of student and teacher engagement in the classroom, and how the resulting analysis could further the role of the teacher in cultivating proof.

The Stylanides (2007) project consisted of data collected from various physical and video archives from 1 third-grade teacher and 22 students for an entire school year. He studied video and audiotape lecture analysis, field notes, transcripts, and copies of students' work. The relationship between the rigor associated with the mathematical argument and an instructional intervention of the teacher was of particular interest. Stylianides concluded that teachers must have a robust knowledge of proof to avoid accepting empirical arguments as proof. Moreover, teachers must assume an active, rather than a passive, role in managing students' proving activity if students hope to develop their proof-writing abilities.

The emphasis of proof amongst elementary students is vastly different than the emphasis for college students, however. The focus shifts from developing a basic understanding of mathematics for young students to demonstrating a deep understanding of mathematics for college students. As such, the instructor must have a genuine grasp of the task at hand.

In 2012, Fukawa-Connelly explored proof through the lens of the student. The author charted a university professor's lecture-based teaching of proof in an undergraduate abstract algebra course. This study's primary focus was the way the instructor presented proof to students since lecture-based mathematics has been widely criticized for proof-based courses (Burton, 1998; Dreyfus, 1999; Fischbein, 1987) and has been blamed for students leaving the field of study altogether (Seymour & Hewitt, 1996).

For the study, Fukawa-Connelly (2012) chose an expert in the field of mathematics with seasoned pedagogical methods and a highly focused objective aimed at developing students' proof-writing abilities. The textbook topics included rings, fields, and group theory with a class of third-year undergraduate students with a calculus background and at least one introductory proof-writing course completed. The research subjects' methods were evaluated using a framework for proof writing developed by Selden and Selden (2003) and Alcock (2010) on methodical approaches to proof writing.

In the article, Fukaway-Connelly (2012) described the various facets of proof writing and the reasoning structure the instructor modeled over 18 class meetings. During the observed classes, 29 proofs were constructed. Of the 29 proofs, 21 were constructed as the instructor asked questions resulting in student feedback which ultimately produced the proof. Student involvement was a large part of the proof presentation.

According to Fukaway-Connelly (2012), this professor's instructional sequence offered students a different learning environment. The instructor consistently modeled, taught and emphasized the modes of thinking necessary in proof writing. Moreover, the professor's lecturing approach underscored significant "aspects of proof framework, hierarchical structure, and the formal-rhetorical parts of a proof," presenting a different student-learning opportunity (Fukaway-Connelly, 2012, p. 342).

CHAPTER III

METHODOLOGY

To investigate conjectural differences between computer-science majors' and mathematics majors' perceptions of proof, data were collected from two Discrete Mathematics classes, one in spring 2015 and the other in spring 2016. Discrete Mathematics is a junior-level introductory proof-writing course and is a degree-plan requirement for both computer science and mathematics degrees at the university under study (See Tables 1 and 2 for required mathematics courses for computer science and mathematics majors).

Table 1

Required Mathematics Courses for Bachelor of Science in Computer Science

Required Course Type	*Course Description*
MATH	Calculus 1
MATH	Discrete Mathematics
MATH	Matrix Methods
OR MATH	Probability and Statistics

Table 2

Required Mathematics Courses for Bachelor of Science in Mathematics

Required Course Type	*Course Description*
MATH	Calculus I
MATH	Calculus II
MATH	Discrete Mathematics
MATH	Abstract Algebra
MATH	Linear Algebra
OR MATH	Matrix Methods
MATH	Elementary Number Theory
MATH	Calculus III
MATH	Differential Equations
MATH	Probability and Statistics

The classes observed were taught by the same professor, using the textbook entitled *Discrete Mathematics* (2009) by Johnsonbaugh. A combined total of 62 students were observed, 35 mathematics majors and 27 computer science majors. The following direct proof was selected, collected from students' exams, and examined for mistakes:

Prove/Disprove: For every integer, a, if a is even then $a + 3$ is odd.

Previous research suggests that students present a wide range of reasoning errors when attempting proof. Some mistakes reflect underlying misconceptions while others have technical imperfections or are otherwise wrong. The grading rubric shown in Table 3 is demonstrative of the views of researchers Annie and John Selden, and was used as a guide for identifying and explaining possible underlying misconceptions and to generalize ideas behind the errors (Selden & Selden, 1987). The rubric is a modified version of the original work of Selden and Selden (1987) and the modified adaption of that rubric from Wheeler and Champion (2011).

Each level of the proof was then examined by the author for problems, compared with the rubric and coded with the appropriate misconception(s) or error(s).

Table 3

Misconceptions and Errors in Proof Codes. (Selden & Selden, 1987; Wheeler & Champion,2011).

Code	*Misconception/Error*	Example
M1	Begin with conclusion	Assume $a + 3$ is odd at the beginning of the proof.
M5	Confuse real number laws	Conclude $\frac{b}{2}$ is no in the integers for b in the integers.
M8	Interfering knowledge	Model a direct proof after an onto proof.
M9	Proof by example	Conclude $a + 3$ by substituting integers.
E1	Misuse of symbols	Use $=$ in place of $\Rightarrow$.
E2	Weak statement	Use a statement stronger than the hypothesis to prove a weaker statement.
E4	Misuse given information	No part of the proof is understandable.
E5	Circularity	Reason a statement back to itself.
E6	Unintelligible proof	No part of the proof is understandable.
E7	Unjustifiable substitution	Replacing defined variables with undefined variables.
E8	Ignore quantifiers	Stating a proof holds for all integers when it only holds for odds.
E9	Logical Holes	Omit multiple consecutive steps in a logical argument.
E11	Computational errors	Mathematical errors (usually algebraic).
E12	Undefined variables	Fail to or incorrectly define variables in the proof.
E13	Set membership	Fail to verify an inverse element is in the domain.
E14	Otherwise incorrect	Wrong or meaningless statement.
NR	No response	No attempt

Variables

Dependent Variable

The dependent variable in this study is the declared major of each participant: computer science or mathematics. Since the dependent variable is categorical with only two possible values, binary logistic regression is the appropriate modeling approach. Statistical Packages for the Social Sciences (SPSS) was used to run analysis of the data. The dependent variable was entered into SPSS as a dichotomous variable and assigned a nominal value.

Table 4
Dependent Variable Encoding

Original Value	Internal Value
CS	0
Math	1

Independent Variables

The continuous independent variables in the study are the 17 misconceptions and errors typically found in students' proofs and previously identified in the proof-coding rubric in Table 3. Each independent predictor variable was entered into SPSS as a quantitative nominal variable. Errors were recorded for E6: unintelligible proof, E11: computational errors, and M9: proof by example; Table 5. The other variables had no influence or no significance on the model therefore were removed.

Table 5

Independent Variables

Independent Variables	Description
M9c	Proof by example
E6c	Unintelligible proof
E11c	Computational errors

Logistic Regression

Since the dependent variable in this study is dichotomous, analysis using the standard regression model, $\mu_y = \beta_0 + \beta_{1x}$, could produce extreme values that fall outside the possible range of values, $p, 0 \leq x \leq 1$, and therefore is not a good fit, (Hosmer, Lemeshow, & Sturdivant, 2013; Moore, McCabe & Craig, 2006). Logistic regression is a method of modeling the relationship between a binary categorical response variable and one or more continuous or categorical explanatory variables. The model is used to analyze the explanatory variables and predict the categorical outcome.

The logistic regression model employs a ratio of proportions for the possible outcomes, referred to as odds:

$$\text{Odds} = \frac{p}{1-p}$$

where p represents the proportion of one outcome, and $1 - p$ represents the proportion of the remaining outcome.

Logistic regression then takes the natural logarithm of the odds proportion

$$\text{Log odds or logit} = \log\left(\frac{p}{1-p}\right) = B_0 + B_1 x$$

which confines possible outcome values, and then the inverse log odds or logit function

$$logit^{-1}x = \frac{e^x}{1+e^x} = \frac{1}{1+e^{-x}}$$

converts the output to values within the appropriate range, $0 \leq x \leq 1$.

Finally, logistic regression is modeled as:

$$p(y_i = 1) = logit^{-1}(X_i B)$$

where p represents the probability, y_i represents the ith observed value for the response variable and $(X_i B)$ represents the linear predictor (Menard, 2010; Moore, McCabe & Craig, 2006).

Hypothesis

The purpose of this study is to determine if the discipline-unique modes of thought associated with participants whose primary field of study is either computer science or mathematics impacts reasoning abilities associated with proof construction.

H_0: Personality profiles associated with college major (computer science or mathematics) have no impact on undergraduate students' proof-writing abilities.

H_1: Personality profiles associated with college major (computer science or mathematics) impacts undergraduate students' proof-writing abilities.

This research will compare proof-writing reasoning performance of students with a declared major of either computer science or mathematics in two Discrete Mathematics classes. Previous studies suggest that computer scientists cling to an algorithmic approach to information (Knuth, 1985) which may limit creative processes required in proof construction (Selden & Selden, 1987; Wheeler & Champion, 2011). Therefore, the expected results of this study were that participants with a background in mathematics would outperform participants with a background in computer science.

CHAPTER IV

RESULTS

Chapter IV presents descriptive statistics and binary logistic regression analysis of the data, including descriptive and output tables. The data were analyzed using Statistical Packages for the Social Sciences (SPSS).

Descriptive Statistics

Descriptive statistics of the dependent variable is presented in Table 6 and descriptive statistics of predictor variables are presented in Table 7. Study participants were categorized by their primary field of study: 43.5% computer science majors and 56.4% mathematics majors. Table 7 shows the number of mistakes and errors participants demonstrated when constructing a direct proof.

Table 6

Descriptive Statistics of Participants

Groups	*N*	Percent
Computer Science Majors	27	43.5
Mathematics Majors	35	56.4

Table 7

Descriptive Statistics of Predictor Variables

Predictors	*N mistakes or errors*	Percent
M9: Proof by example		
Computer Science	3	4.8
Mathematics	0	0
E6: Unintelligible proof		
Computer Science	6	9.7
Mathematics	2	3.2
E11: Computational error		
Computer Science	7	11.3
Mathematics	4	6.5
Total mistakes and errors		
Computer Science	16	25.8
Mathematics	6	9.7

Logistic Regression Analysis

A successful logistic regression model allows for prediction of an outcome based on various predictor variables. The data output will indicate if the predictor variable(s) have a statistically significant effect on the prediction, and if so, how well the model predicts the outcome.

The first tables produced by SPSS are the Block 0: Beginning Block (Table 8), Variables in the Equation (Table 9), and Variables not in the Equation (Table 10). The

Block 0 output excludes predictor variables and assumes all participants are math majors. This model, sometimes called the intercept model or null model, establishes a baseline, or point of reference, by which to compare the effectiveness of the model once predictor variables are introduced. If this was the only model available, the assumption would correctly classify $\frac{35}{62} \approx 56.5\%$ percent of all cases.

Table 8

Block 0: Beginning Block

Classification Table[a,b]

			Predicted		
			groups		
	Observed		CS	Math	Percentage Correct
Step 0	groups	CS	0	27	.0
		Math	0	35	100.0
	Overall Percentage				56.5

a. Constant is included in the model. b. The cut value is .500

The predicted odds of success, represented in Table 9, column Exp(B), is $\frac{35}{27} \approx 1.296$. In the absence of predictor variables, the model has no statistical significance at $p = .311$ with 1 degree of freedom.

Table 9

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	.260	.256	1.026	1	.311	1.296

Table 10 presents the predictor variables that will be introduced in the SPSS secondary output, Block 1: Method = Forward Stepwise (Likelihood Ratio). The variables M9c, E6c, and E11c are statistically significant at $p < .05$ with 1 degree of freedom.

Table 10

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	M9c	4.087	1	.043
		E6c	5.018	1	.025
		E11c	4.413	1	.036
	Overall Statistics		15.540	3	.001

With the introduction of predictor variables, the model takes a more compelling shape. Block 1 output begins with the Omnibus Tests of Model Coefficients, a likelihood ratio chi-square test of the current model's statistical significance versus that of the intercept or null model (IBM, 2013). Each step indicates the addition of a new predictor variable. SPSS includes the variable E6c in step 1, E6c and M9c in step 2, and E6c, M9c, and E11c in step 3.

The results shown in Table 11 reveal a chi-square with 1 degree of freedom, $\chi^2(1) = 5.128$, and a significance value of $p = .024$ in step 1; a chi-square with 2 degrees of freedom, $\chi^2(2) = 11.277$, and a significance value of $p = .004$ in step 2; and finally a chi-square with 3 degrees of freedom, $\chi^2(3) = 17.435$, and a significance value of $p = .001$ in step 3.

The step and block rows compare the -2 log likelihood of the newest model to the previous version to determine if the new variables are causing improvement. In each step, the chi-square test statistic produces a significance-level decrease which suggests that each new model, with the addition of a new predictor variable, more accurately predicts the outcome than the previous model. While each model is better than the last, the model in step 3 with all three independent variables has the greatest degree of statistical significance. Therefore, the decision is to reject the null hypothesis.

Table 11

Block 1: Method Enter = Forward Stepwise (Likelihood Ratio)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	5.128	1	.024
	Block	5.128	1	.024
	Model	5.128	1	.024
Step 2	Step	6.149	1	.013
	Block	11.277	2	.004
	Model	11.277	2	.004
Step 3	Step	6.158	1	.013
	Block	17.435	3	.001
	Model	17.435	3	.001

The Model Summary, shown in Table 12, provides the -2 log likelihood, an approximation of how poorly the model predicts success, and the pseudo-R^2 values which approximate how much variation in the predicted outcome is explained by the new model. This study's model is represented in step 3 in Tables 12-16.

The -2 log likelihood is 67.480 in step 3, a decrease from the previous steps. The desired result with the -2 log likelihood is small statistic, and the decrease from step 1 to step 3 indicates improvement in this study's model. The pseudo-R^2values, Cox & Snell R^2 = .245 and Nagelkere R^2= .329, indicates that the predictor variables explain between 24.5% and 32.9% of the variation and suggest a moderate to strong correlation between the variables.

Table 12

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	79.787[a]	.079	.106
2	73.638[b]	.166	.223
3	67.480[b]	.245	.329

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001. b. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

The Hosmer and Lemeshow Test, shown in Table 13, assesses the model's goodness of fit, or how accurately the data fits the model. The desired result with the Hosmer-Lemeshow Test is a large statistic, $p > .05$. The step 3 model results show a Hosmer-Lemeshow chi-square with 2 degrees of freedom, $\chi^2(2) = .150$, and the results are not statistically significant at $p = .928$ which suggests the model is a good fit.

Table 13

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.000	0	.
2	.000	1	1.000
3	.150	2	.928

The Classification Table, Table 14, represents the observed and predicted classification for the model. Binomial logistic regression estimates the probability of an event occurring based on specific parameters established within the model. In this study, the probability of a participant being a mathematics major is based on the occurrence or non-occurrence of two specific errors and one misconception within the proof. If the estimated probability is greater than 50%, SPSS classifies the participant as a math major.

The SPSS baseline or null output correctly predicted 56.5%. Once predictor variables were added to the model, the proportion of cases correctly predicted increased to $\frac{17}{27} \approx 63.0$ % for computer science majors, and $\frac{29}{35} \approx 82.9$ % for mathematics majors.

Overall, the model correctly predicted $\frac{46}{62} \approx 74.2\ \%$ of the cases, an increase of 31% over the null model.

Table 14

Classification Table[a]

			Predicted		
			groups		
	Observed		CS	Math	Percentage Correct
Step 1	groups	CS	7	20	25.9
		Math	2	33	94.3
	Overall Percentage				64.5
Step 2	groups	CS	10	17	37.0
		Math	2	33	94.3
	Overall Percentage				69.4
Step 3	groups	CS	17	10	63.0
		Math	6	29	82.9
	Overall Percentage				74.2

a. The cut value is .500

The impact each variable has on the model is seen in Table 15, Variables in the Equation. This table shows the joint association between each variable through the regression coefficient (B), the Wald statistic, and the odds ratio. The Wald statistic determines statistical significance between the variables. The variables E6c and E11c added significantly to the model prediction with p=.019 and p=.017 respectively. While

M9c, with p=.999, did not add statistical significance by itself, it contributed to the overall correlation and helped improve the overall predictive power of the model.

The B coefficients are useful for predicting the probability of an event occurring. This statistic represents the log odds for each one-unit change in an independent variable when all others are constant. All three predictive variables have a negative B value which suggests the likelihood of a mathematics student making the mistake is low.

The odds ratio, represented by Exp (B), expresses the change in odds for each one-unit increase of the independent variable. The variable E6c odds ratio is Exp(B) = .124, however, the inverted odds ratio, $\frac{1}{\mathrm{Exp(B)}} = \frac{1}{.124} \approx 8.06$ may be easier to interpret. The inverted ratio suggests that for each decrease in one unit of this variable, the odds of the participant being a mathematics major increases by a factor of 8.06. In other words, a mathematics major is 8.06 times less likely to make this mistake than a computer science major. The variable E11 has an odds ratio Exp(B) = .184 and an inverted odd ratio $\frac{1}{.184} \approx 5.43$ which indicates that a mathematics major is 5.43 times less likely to make this mistake than a computer science major.

Table 15

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	E6c	-1.754	.850	4.252	1	.039	.173
	Constant	.501	.283	3.123	1	.077	1.650
Step 2[b]	M9c	-21.866	23205.422	.000	1	.999	.000
	E6c	-1.916	.856	5.016	1	.025	.147
	Constant	.663	.299	4.936	1	.026	1.941
Step 3[c]	M9c	-22.285	23205.422	.000	1	.999	.000
	E6c	-2.088	.887	5.537	1	.019	.124
	E11c	-1.691	.710	5.667	1	.017	.184
	Constant	1.082	.366	8.751	1	.003	2.950

a. Variable(s) entered on step 1: E6c.

b. Variable(s) entered on step 2: M9c.

c. Variable(s) entered on step 3: E11c.

Hypothesis Results

H_0: Personality profiles associated with computer science or mathematics have no impact on undergraduate students' proof-writing abilities.

H_1: Personality profiles associated with computer science or mathematics impacts undergraduate students' proof-writing abilities.

A binomial logistic regression was performed to ascertain the effects of algorithmic and mathematical modes of thinking on the likelihood that participants are collegiate mathematics majors. The logistic regression model was statistically significant, $\chi^2(3) = 17.435$, $p = .001$. The model explained 32.9 % (Nagelkerke R^2) of the variance and the overall positive-predictive value was 74.9%. The model correctly predicted 63% of computer science majors and 82.9% of mathematics majors.

Of the three predictor variables, two were statistically significant: E6c, unintelligible proof and E11c, computational errors. Mathematics majors were 8.06 times less likely to write an unintelligible proof and 5.43 times less like to make computational errors in proof than computer science majors. Based on this analysis, the model successfully discriminated between the two disciplines and therefore established a connection between the modes of thinking and ability to construct proof. The decision is to reject the null hypothesis and accept the alternative.

CHAPTER V

DISCUSSION

In this study, mistakes and errors found in undergraduate students' proofs were used to predict how students traditionally associated with an algorithmic or mathematical approach to thinking would perform when constructing proofs. Current literature has documented a wide range of difficulties associated with learning to write proof, but none have linked modes of thinking to the process. This idea is significant since it has been established that most undergraduate-mathematics courses are substantially algorithmic which contributes to an algorithmic approach to knowledge.

Despite the small number of possible mistakes and errors considered in this study, the results are revealing. This study is based on three independent predictor variables: one potential mistake, and two potential errors participants make when attempting to write a direct proof. Of the three predictors, two are serious and suggest the student has a poor understanding of proof.

The severe error is logically unintelligible proof (E6). This type of error presents itself as a series of mathematical statements and symbols, however, neither the individual statements nor the proof in its entirety can be understood. Moreover, the statements are unintelligible or incorrect. The severe mistake is proof by example, (M9). This type of mistake presents itself as a demonstration of the general statement and implies that because the statement is true for one case, it is true for all cases.

Of the proofs written by computer science participants in this study, one in three contained a severe mistake or error, whereas mathematics majors' proofs contained one serious mistake or error for every 17.5 proofs constructed (Table 7). This represents a substantial difference between computer science majors' and mathematics majors' demonstration of proof-writing competence.

Limitations and Future Research

Mathematics History

The mathematics background of each student was not included in this study. Aside from a prerequisite for Calculus 1, no other college mathematics history is assumed. It is possible the students had a diverse range of exposure to formal proof prior to Discrete Mathematics, which could impact the outcomes. Future studies could include each students' mathematics history from high school- to college-level courses. Moreover, students' previous exposure to proof should be measured and considered a significant contributing factor.

Sample

For this study, a single direct proof from the second exam was collected, coded, and analyzed. Proofs taught in this Discrete Mathematics class ranged from mathematics induction, one-to-one proofs, onto proofs, divisibility proofs, as well as other direct and indirect proofs. It is possible that exploring a larger sample containing a variety of proofs could produce different results.

Another sample concern is that participants were not randomly chosen; the sample was collected from two Discrete Mathematics classes from two semesters. The reason for this is the sample was taken from a university with a small mathematics and computer science department and Discrete Mathematics is only offered once per academic year and sometimes in the summer. It is widely recognized that unfavorable sampling techniques have the potential for bias contamination; therefore, future studies could be designed to incorporate randomization within the sample.

Finally, while the sample for mathematics majors was large enough at $n = 35$, the sample of computer science participants was smaller than desired at $n = 27$. It is possible that a larger sampling of computer science majors could alter the results.

Conclusion

This study produced a good model for predicting a students' approach to knowledge based on mistakes and errors produced in proof construction. Future studies could include mathematics backgrounds of students which could impact the way the proof-writing course is designed, as well as prerequisites for the class.

REFERENCES

Alcock, L. (2010). Mathematicians' perspectives on the teaching and learning of proof. IN: Hitt, F., Holton, D. and Thompson, P. (eds). *Research in Collegiate Mathematics Education* (VII, pp. 63-92).

Burton, L. (1998). The practices of mathematicians: What do they tell us about coming to know mathematics? *Educational Studies in Mathematics, 37*(2), 121-143. Retrieved from http://ezproxy.twu.edu:2069/stable/3483312

de Villiers, M. (1991, November). The role and function of proof in mathematics. *Pythagoras*, 17-24. Retrieved from Academic Search Complete.

Dreyfus, T. (1999). Why Johnny Can't Prove. *Educational Studies in Mathematics, 38*(1/3), 85-109.

Fischbein, E. (1987). *Intuition in science and mathematics: An educational approach* (Vol. 5). N.p.: Dordecht.

Fukawa-Connelly, T. P. (2012, November). A case study of one instructor's lecture-based teaching of proof in abstract algebra: making sense of her pedagogical move. *Educational Studies in Mathematics*, *81*(3), 325-345.

Gokkurt, B., Soylu, Y., & Sahin, O. (2014, December). Analysis of the mathematical proof skills of students of science teaching. *Educational Research Quarterly*, *38*(2), 3-22.

Gries, D., & Schneider, F. B. (1995, October). Teaching math more effectively, through calculation proofs.*The American Mathematical Monthly*, *102*(8), 691-697.

Griffiths, P. (2000). Mathematics at the Turn of the Millennium. *The American Mathematical Monthly*, *107*(1), 1-14. doi:1. Retrieved from http://ezproxy.twu.edu:2069/stable/2589372 doi:1

Hanna, G. (2000). Proof, explanation and exploration: an overview. *Educational Studies in Mathematics*, *44*(1/2), 5-23.

Hosmer, D. W., Jr., Lemeshow, S., & Sturdivant, R.X. (2013) *Applied Logistic Regression* (3rd ed.). Hoboken, NJ: Wiley.

IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

Inglis, M., & Alcock, L. (2012, July). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, *43*(4), 358-390.

Johnsonbaugh, R. (2009). *Discrete Mathematics* (7th ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Kitcher, P. (1984). *The nature of mathematical knowledge*. New York, NY: Oxford University Press.

Knuth, D. E. (1985, March). Algorithmic thinking and mathematical thinking. *The American Mathematical Monthly*, *92*(3), 170-181.

Knuth, D. E. (1974, April). Computer science and its relation to mathematics. *The American Mathematical Monthly*, *81*(4), 323-343.

Menard, S. (2010). *Logistic regression: From introductory to advanced concepts and applications*. London: SAGE Publications.

Moore, D. S., McCabe, G. P., & Craig, B. A. (2006). *Introduction to the Practice of Statistics* (8th ed., pp. 14-1-14-27). New York, NY: W. H. Freeman and Company. Retrieved from http://bcs.whfreeman.com/webpub/statistics/ips8e/student%20resources/companion%20chapters/c14LogisticRegression.pdf

National Council of Teachers of Mathematics [NCTM]. Principles and standards for school mathematics (2000). In *National Council of Teachers of Mathematics*. Retrieved September 26, 2016.

Rav, Y. (1999). Why do we prove theorems. *Philosophia Mathematica*, *7*(3), 5-41.

Selden, A., & Selden, J. (1987, July). Errors and misconceptions in college level theorem proving [Electronic version]. *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, III*, 457-470.

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, *34*(1), 4-36.

Seymour, E., & Hewitt, N. M. (1996). *Talking about leaving: why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Stylianides, A. J. (2007, May). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, *38*(3), 289-321.

Uhlig, F. (2002). The role of proof in comprehending and teaching elementary linear algebra. *Educational Studies in Mathematics*, *50*(3), 335-346.

Weber, K. (2008). How mathematicians determine in an argument is a valid proof. *Journal for Research in Mathematics Education*, *39*(4), 431-459.

Weber, K., & Mejia-Ramos, J. P. (2013). Mathematics majors' beliefs about proof reading. *International Journal of Mathematical Education in Science and Technology*, *45*(1), 89-103.

Wheeler, A., & Champion, J. (2011). Students' proofs of one-to-one and onto properties in introductory abstract algebra. *International Journal of Mathematical Education in Science and Technology*, *44*(8), 1107-1116.