# Relationship Between Students' Proof Schemes and Definitions

David Bryant Plaxco

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Mathematics

Anderson H. Norton, Committee Chair
Nicholas A. Loehr
Jesse L. M. Wilkins

May 4, 2011
Blacksburg, VA

Keywords: proof, definition, probabilistic independence, mutual exclusivity, mathematics education, undergraduate education

Relationship Between Students' Proof Schemes and Definitions

David Bryant Plaxco

ABSTRACT

This research investigates relationships between undergraduate students' understanding of proof and how this understanding relates to their conceptions of mathematical definitions. Three students in an introductory proofs course were each interviewed three times in order to assess their proof schemes, understand how they think of specific mathematical concepts, and investigate how the students prove relationships between these concepts. This research used theoretical frameworks from both proof and definition literature. Findings show that students' ability or inability to adapt their concept images of the mathematical concepts enhanced and impeded their proof schemes, respectively.

## Acknowledgements

I would like to thank my committee — Doctors Norton, Loehr, and Wilkins — for their invaluable support and scrutiny of this research.

I would like to thank my family and friends who supported and encouraged my efforts: Mom, Dad, Jeri Sue, Michael, Anna, Mitchell, Jessica, Johnnie, Nabil, Adyan, Emily, and Tyler.

I would also like to thank my many friends and classmates in both the Mathematics Department and School of Education: Andy, Matt, Steven, Angie, Mike, Alexis, Jean, Beth, Joyce, and any other unlucky person who would listen to my rants about probability.

## Table of Contents

# Introduction

## Researching Proof

In recent years, undergraduate mathematics proofs courses have gained increased attention for curriculum improvement (Alcock & Simpson, 2002; Harel & Sowder, 2007; Ko, 2010). This emphasis on the undergraduate courses is partially an effect of the emphasis on proof over the entire American K-12 mathematics curriculum (National Council of Teachers of Mathematics [NCTM], 2000; Ko, 2010) and, in turn, the need to educate teachers and preservice teachers about the importance of proof in learning any and all mathematics (Martin & Harel, 1989; Knuth, 2002).

Proofs courses mark a vital transition in the undergraduate curriculum. The knowledge and skills acquired in a proofs course can empower students to more capably verify or refute new mathematical ideas that they encounter. Before taking such courses, students are typically expected to apply known theorems to specific situations without general proofs of the theorems themselves. Higher-level undergraduate and graduate courses in mathematics generally have high expectations of students' abilities to prove conjectures and theorems (Alcock & Simpson, 2002). Current research, however, finds many undergraduate students' proving skills to be lacking even after a proofs course (Martin & Harel, 1989). In order to improve students' ability to prove mathematical statements, we must better understand how students approach mathematical proof.

Under the prevailing undergraduate curriculum, we can think of a proofs course as the point in a mathematics student's undergraduate experience where they transition from "mathematics student" to "mathematician" (Alcock & Simpson, 2002; Alibert & Thomas, 2002). One must acknowledge that distinguishing between these two terms has serious pedagogical implications (Harel & Sowder, 2007; NCTM, 2000; Nickerson & Rasmussen, 2009), which is why this is discussed in the context of the *prevailing* undergraduate curriculum. Highlighting this transition points to the need for improvement in how we teach proofs courses. Even when a student's curriculum has provided rich experiences that have allowed the student to develop more sophisticated logico-mathematical skills, this student may still lack some of the "tools" acquired in a proofs course (e.g. the Principle of Mathematical Induction, higher-level mathematical notation).

A more basic argument toward the importance of a proofs course (and not only for mathematics majors) considers the basic act of knowing and learning. We can think of knowledge with a constructivist approach to learning, typical of Piaget (1970) or von Glasersfeld (1995). In this mindset, knowledge is our general ability to predict the effects of operations in a given system. When one observes actions and results in a system, one must: 1) somehow convince oneself that the results are a logical conclusion of how one understands the system, 2) modify one's understanding to include a new operation or scheme so that this new information complies with previous understanding, or 3) reject the new information as invalid, allowing one's old operations to continue unscathed. This progression describes the fundamental process of building knowledge. Proofs courses should provide undergraduates with a rich opportunity to practice the first outcome of this process in a mathematical and logical way. Since a large part of mathematical proof is convincing oneself of the truth of an argument or relationship, this first outcome is all too critical (Harel & Sowder, 1998).

The heightened emphasis throughout the literature on how the concept of mathematical proof is addressed throughout the entire United States mathematics curriculum supports the need for this research (Ko, 2010; NCTM, 2000). The general purpose of this investigation is toward a better understanding of “what are students’ current conceptions of proof,” “what are students’ difficulties with proof,” and “what accounts for these difficulties” (Harel & Sowder, 2007, p. 3). This research will attempt to address these questions, thus informing the undergraduate curriculum. From this, college instructors can improve undergraduate proofs courses and influence preservice teachers’ disposition of the role of reasoning proof in the elementary and secondary classroom.

Several researchers have noted the seeming inseparability between mathematics and proof (Harel & Sowder, 1998; Jaffe & Quinn, 1993; Schoenfield, 1994; Wu 1996). The three main areas in need of further research toward the understanding of students’ proof schemes and how to teach proof are as follows: 1) epistemological analysis, 2) empirical research, and 3) design research (Ball, Hoyles, Jhanke, & Movshovitz-Hadar, 2002). Respectively, these three areas address the following domains of knowledge:

1. Increasing teachers’ understanding of current, higher-level mathematical proofs and how mathematicians approach them,
2. Understanding the process through which students learn proof,
3. Making our curricula better suited for teaching proof.

This study will fall under the second motivation, empirical research, with the purpose of investigating students’ use of definitions in proofs and how the mathematical definitions are used to attain understanding of relationships in mathematics.

**Researching Definition**

Definition and proof are virtually inseparable. Mathematical definitions reflect the rigor and precision on which mathematics is founded. They must be clear, concise, and minimal (Vinner, 1991). It should be noted also that these adjectives are equally applicable to traditional mathematical notions of proof (Harel & Sowder, 2007). Even within the same system, if two people have differing definitions of a single concept, basic axioms can fail. For example, consider proving the limit as $x \to \infty$ of the function of real numbers $f(x) = x$. One person may think the real numbers are the set $(-\infty, \infty)$, whereas another person could include $\pm\infty$ in the real numbers. To one person the limit diverges, while to the other person it converges to $\infty$. While the distinction in this case may be quick to resolve, there are many more cases where a common ground would not be as easy to find. This is evident in the more infamous mathematical debates throughout history (e.g., existence of infinitesimals or "fluxions").

We can think generally about how definitions relate to proof. It seems that, given that one has a fairly complete understanding of a concept's formal definition, one may or may not be able to prove various relationships involving said definition. It seems less likely that a person could have an incomplete or incorrect understanding of the definition of a mathematical idea and formally prove a relationship involving that concept. But what is the exact nature of the relationship between these two aspects of mathematics? Surely it is possible that a student can understand advanced mathematical techniques for proving, understand all of the concepts being investigated independent of each other, and still fail to prove a relationship between them. Could a student logically convince himself and others of a relationship using an incorrect definition? Furthermore, could a student convince himself and others of a relationship using incorrect techniques and an incorrect definition?

**Finding a Context for Comparison**

During a graduate course that explored students' understanding of probability and statistics, I came across research that investigated undergraduates' understanding of the relationships between independent and mutually exclusive events. Manage and Scariano (2010) found that most of the students in their study thought that two events being independent implied that they were mutually exclusive and vice versa. One's initial reaction may be to conclude this exact relationship. It is only after careful consideration of the two concepts' meanings that one realizes the two terms have almost exactly opposite meanings. We see that this could provide an ideal context for exploring the use of definition in proof.

One aspect of this context that is so appealing is that typically, after learning the concepts of independence and mutual exclusivity in middle or high school, students seldom encounter formal situations in which they must describe events as either mutually exclusive or independent. Because of this, we often neglect the formal definitions of what it means to be mutually exclusive or independent. People tend, however, to take for granted that some relationships are independent. This phenomenon is common, for instance, when someone is asked to calculate the probability of a compound event. Most people believe that multiplying two probabilities yields a correct calculation of the intersection's theoretical probability, but one could argue that these people seldom think about why they multiplied the two probabilities.

It can be argued that in order for the high number of students in Manage and Scariano's study to assert an incorrect relationship between mutually exclusive and independent events, they must have convinced themselves of this misguided relationship. This process can be thought of as their own *internal conviction* that the concepts of mutual exclusivity and independence are almost synonymous (Harel & Sowder, 1998). Since this conception was so prevalent among students in a probability and statistics course (where the formal mathematical definition of each was most likely presented), it seems to follow that it could be at least as prevalent among mathematics majors who have not had an undergraduate probability and statistics course. It follows, then, that students in a proofs course should have varying proof schemes and degrees of understanding of the formal definitions of mutual exclusivity and independence. We can then observe how these students might convince themselves of such a relationship.

**Purpose**

**The purpose of this research is to investigate students' conceptions of mathematical definitions and how these relate to their general proof schemes.** This research intends to do this through a three-step process with the following objectives:

1) to assess and categorize undergraduate students' general proof schemes,
2) to observe undergraduate students' use and understanding of the mathematical definitions of mutually exclusive and independent events,
3) to observe and better understand the process of a student attaining proof of a mathematical concept by
    - asserting their perceived relationship of mutually exclusive and independent events
    - attempting to convince someone of their claim.

These objectives will be accomplished through an examination of specific examples of how students' proof schemes relate to their understanding of the mathematical definitions of the concepts involved. I will conduct a series of three interviews with undergraduate students so that each interview is directed toward the three purposes listed above. After assessing students' proof schemes and observing their conceptions of independence, I will relate and reconcile these two aspects of the students' mathematical understanding by observing their proof of an asserted relationship.

# Literature Review

## An Approach to Proof Schemes

Harel and Sowder (1998, 2007) have provided a fundamental framework for research in students' conceptions of proof. This framework begins by defining proof as a two-part process: ascertaining and persuading. Ascertaining occurs when a student convinces him or herself that a concept is true. Attainment of truth for oneself does not necessarily mean that this self-convincing process is logical, deductive, or even mathematical. It merely means that that individual believes the relationships under discussion to be true. Persuading is a more complicated process in that the student must then communicate his or her conception of the relationship to another person in a compelling way. The argument used to convince the other person may not necessarily be the same argument that convinced the first person.

A person's notion "of what constitutes ascertaining and persuading" is called that person's proof scheme (Harel & Sowder, 1998, p. 244). There are varying levels of proof schemes described in the literature. Broadly, these proof schemes are External, Empirical, and Analytical — an External proof scheme includes validations by authority, ritual, or symbolism; an Empirical proof scheme is inductive or perceptual; and an Analytical proof scheme is more rigorous and logical (Harel & Sowder, 1998). It is important to note that, at a given time, a person may not completely display evidence of exactly one proof scheme. Because of this, students' proof schemes are generalizations of the types of proof schemes evident through their work. For example, a student could exhibit both Empirical and Analytical proof schemes within a short period of time or even within a single proof; such a student could be said to have an "emerging Analytical" or "Empirical/Analytical" proof scheme.

Many researchers have built upon Harel and Sowder's framework. For example, Housman & Porter (2003) investigated the proof schemes of "above average" students. Their results confirm that some students can exhibit all three major types of proof scheme over a short amount of time. One interesting result from their work is that none of their research participants exhibited both analytical and external conviction proof schemes without exhibiting empirical proof schemes.

Other important additions to Harel and Sowder's framework come from Knuth (2002) and Weber (2001, 2009) among others (Selden & Selden, 2003; etc.). While Knuth focuses mainly on teachers' beliefs about proof, Weber is more concerned with undergraduates' understanding as future mathematicians. Weber (2009) investigated how students read proofs and the types of proofs students find convincing. Interestingly, Weber (2009) found that none of the students he interviewed (who had recently completed an introduction to proofs course) displayed an empirical proof scheme. He supports this research by outlining in his earlier research the typical classroom expectations of mathematics professors, which tend to rely on students copying and studying proofs presented to them in class (Weber, 2004). Thurston (1994) also addresses this Definition-Theorem-Proof (DTP) approach to teaching graduate- and undergraduate-level mathematics courses. Thurston is quick to note the deficient nature of DTP arguing that, "A clear difficulty with the DTP model is that it doesn't explain the source of the questions" (1994, p. 39).

Alcock and Weber (2005) differentiated between students' writing and reading proofs in a study that investigated students' line-by-line analyses of a proof that was invalid because of an unsupported inference. The researchers analyzed thirteen students' discussion of the proof. Only two of these students legitimately explained how the presented proof was flawed according to the authors. Five students accepted the proof. The other six students, three of whom arguing that the formal definition was not presented, found flaws in the proof other than the proof's incorrect inference (Alcock & Weber 2005).

Weber and Alcock (2004) also contribute a framework for discussing students' semantic and syntactic proof productions, which can be described by the students' use of diagrams and instantiations versus strict manipulation of symbols, respectively. While this framework is constructed outside of Harel and Sowder's (1998) proof schemes, the two classification systems for students' mathematical tendencies seem as though they could work successfully together as neither excludes the other.

**An Approach to Definition**

Vinner (1991) distinguishes between the ideas of concept image and concept definition. He defines the *concept image* as a non-verbal entity such as a "visual representation of the concept… [or] a collection of impressions or experiences" (p. 68) which our mind associates with the concept. In contrast, the *concept definition* is the formal mathematical definition of a concept. These two ideas are not necessarily the same thing. Vinner uses the sentence, "my nice green car is parked in front of my house," as an example of concept image (p. 67). He argues that the reader or listener does not necessarily consider the definition of each word in the sentence, but that each word invokes a generic image in their mind, the collection of which allows the sentence to take form as a whole impression.

Relating this to mathematics, we understand that mathematical concepts can take any number of images in a person's mind and seldom the same image among many people (Vinner, 1991). Thurston (1994) provides an excellent example of various conceptions of the derivative:

"The derivative function can be thought of as:

(1) Infinitesimal: the ratio of the infinitesimal change in the value of a function to the infinitesimal change in a function.
(2) Symbolic: the derivative of $x^n$ is $nx^{n-1}$, the derivative of sin(x) is cos(x), the derivative of $f \circ g$ is $f' \circ g * g'$, etc.
(3) Logical: $f'(x) = d$ if and only if for every ε there is a δ such that when $0 < |\Delta x| < \varepsilon^\star$,

$$\left|\frac{f(x+\Delta x)-f(x)}{\Delta x}-d\right|<\delta$$

(4) Geometric: the derivative is the slope of a line tangent to the graph of the function, if the graph has a tangent.
(5) Rate: the instantaneous speed of f(t), when t is time.
(6) Approximation: The derivative of a function is the best linear approximation to the function near a point.
(7) Microscopic: The derivative of a function is the limit of what you get by looking at it under a microscope of higher and higher power." (p. 39-40)

While some of these {2, 3, 5} are symbolic and not graphic (hence they do not strictly follow the idea of a concept image) others {4, 6, 7} are readily accessible images that we can relate to the term "derivative." It can be argued that the first representation of the derivative is both a concept image and a concept definition according to some.

Edwards and Ward (2004) provide insight into students' use and misuse of mathematical definitions in proofs. They base much of their framework on Landau's (2001) distinction between *extracted* and *stipulated* definitions as well as Vinner's (1991) framework. The phrases "extracted" and "stipulated" coincide closely with concept image and concept definition, respectively. Edwards and Ward conclude that, "the special nature of mathematical definitions should be addressed more directly in mathematics courses at all levels, but especially in the first proof-intensive course" (2004, p. 422). This highlights the need for research into the relationship between proof and definition.

We see a striking similarity between the proof literature and definition literature that one could argue is more than merely coincidental. Let us consider semantic and syntactic proof

---

⋆ Author wrote "δ" instead of "ε".

production in the context of concept image and concept definition. Weber and Alcock (2004) found that a student syntactically proving a relationship "would need to be able to recite the definition of a mathematical concept as well as recall important facts and theorems concerning that concept. The prover would also need to be able to derive valid inferences from the concept's definition and associated facts" (p. 229). This description of syntactic proof production seems especially related to Vinner's concept definition. Similarly, Weber and Alcock argue that a student producing a semantic proof would necessarily need to instantiate relevant mathematical objects. Such instantiations would require at least some form of concept image.

**Independence and Mutual Exclusivity**

Probability and statistics have been increasingly emphasized in elementary and high school education over the past two decades (NCTM, 1989; NCTM, 2000). This emphasis has also extended into post-secondary education. Mutual exclusivity and independence have gained such emphasis, making it important to accurately gauge students' conceptualizations of these two ideas (Manage & Scariano, 2010; D'Amelio, 2009). Manage and Scariano (2010) found that an alarmingly high percentage of undergraduate students who were enrolled in a course in probability and statistics had fundamental misunderstandings about the relationship between the ideas of independent events and mutually exclusive events. They found this by directly assessing students' understanding of this relationship through a non-scientific, multiple-choice survey of 217 students.

Manage and Scariano (2010) explored student responses to two questions. In each question, two events, A and B, are assumed to have nonzero probabilities in the same sample space. The first question presented a Venn diagram of the sample space with non-intersecting areas, labeled "*A*" and "*B*," within a rectangular sample space. The question stated that *A* and *B* are mutually exclusive and gave four responses: "(a) *A* and *B* are independent events," "(b) *A* and *B* are not independent events," "(c) *A* and *B* may or may not be independent events," and "(d) I really don't know how to do this problem" (Manage & Scariano, 2010, p. 18). The second question presented the fact that *A* and *B* are independent, have nonzero probabilities, included no diagram, and gave four responses: "*A* and *B* are mutually exclusive events," "*A* and *B* are not mutually exclusive events," "*A* and *B* may or may not be mutually exclusive events," and "I really don't know how to do this problem" (Manage & Scariano, 2010, p. 19).

In the first question, 68.3% of students incorrectly chose the first answer choice, namely "*A* and *B* are independent events." In total, 88% did not choose the correct answer choice B. With respect to the second question, 36% of students gave the incorrect first response, "*A* and *B* are mutually exclusive events," whereas 23.3% of students responded correctly with answer choice (b). Results from the first question indicate that students seem to think that these two ideas have a direct relationship, that mutual exclusivity implies independence. This misconception seems less prevalent in the second question, since the responses were more evenly distributed than they were in the first (Manage & Scariano, 2010).

D'Amelio (2009) found that most participants could not correctly identify a method for calculating the probability of the union of mutually exclusive events. Most students mistook the product, rather than the sum, of the two events' respective probabilities for the proper calculation. These results suggest students' confusion about the use of such a product when calculating certain probabilities. They also point to a similar misconception to that found in Manage and Scariano (2010), particularly a misunderstanding of the distinction between independent and mutually exclusive events. Confusion between calculating the probability of the intersection of two events rather than their union provides an alternative explanation for this mistake.

Shaughnessey (1992) identifies two equivalent definitions of independent events given that the two events $A$ and $B$ are in the same sample space and have nonzero probabilities (Table 1).

| |
|---|
| **Definition 1-** $P(A\|B) = P(A)$ **Definition 2-** $P(A \cap B) = P(A) \times P(B)$. |

**Table 1. Definitions of Independence**

Both D'Amelio (2009) and Manage and Scariano (2010) use Definition 2 in their research. Each of these researchers also define two events $A$ and $B$ as "*mutually exclusive* if and only if ($A$ and $B$) = $A \cap B = \varnothing$" (Manage, p. 15). From this definition, if $A$ and $B$ are mutually exclusive then $P(A \cap B) = 0$. So, by the zero product property, it cannot be the case that $P(A \cap B) = P(A) \times P(B)$ and $P(A \cap B) = 0$ when $A$ and $B$ have nonzero probabilities.

In the available research, the explicit relationship between independence and mutual exclusivity was found in only three articles (Keeler & Steinhorst, 2001; Kelly & Zwiers, 1988; Manage & Scariano, 2010). Manage and Scariano use the reasoning discussed above, whereas Kelly and Zwiers address this relationship in the context of student misunderstanding. They provide several examples of how each of these ideas can be explored separately in a classroom. Kelly and Zwiers contend that, "most of the confusion arises because we, as instructors, do not take the time to relate the two concepts" (p. 100). They then blatantly state the relationship between the two ideas- "mutually exclusive events are (almost) never independent" (p. 100). They attribute the "almost" in this last quote to the "pathological cases" when one or both events considered have zero probability (Kelly & Zwiers, 1988).

In considering pedagogical implications for this research, we find that students have many difficulties with both conditional probability and independence (Shaughnessey, 1992). This research goes on to say that students' "misconceptions of conditional probability may be closely related to students' understanding of independent events and of randomness in general" (Shaughnessey, 1992, p. 475). He points out that many researchers "advocate introducing the concept of independence via the conditional probability definition (Definition 1), as they believe this is more intuitive for students" (1992, p. 475). This intuition comes in the context of without-replacement problems. If the sample space remains unchanged, then the first experiment bears no effect on the second experiment.

Other pedagogical research in this area discusses students' misconceptions related to independence almost exclusively with respect to conditional probability (Tarr & Lannin, 2005). Tarr and Lannin (2005) justify their concentration on these types of misconceptions, citing Shaughnessey (1992), and focus on replacement and non-replacement situations because of the prevalence of these types of problems in the typical curriculum. Tarr and Lannin state that "within this context [that of *with-replacement* situations], an 'understanding of independence' is demonstrated by students' ability to recognize and correctly explain when the occurrence of one event does not influence the probability of another event" (2005, p. 216). There is also an emphasis that students understand the change of an event's probability in non-replacement conditional probability problems is due to the change of the sample space.

While Shaughnessey (1992), Tarr and Lannin (2005), and Keeler and Steinhorst (2001) suggest conditional probability as a context for independence, D'Amelio (2009), Kelly and Zwiers (1988), and Manage and Scariano (2010) each explore student misconceptions outside of the conditional probability setting. This could provide some reasoning into why D'Amelio

(2009) and Manage and Scariano (2010) had such disturbingly low numbers of correct responses. Supposing that the students' previous curricula addressed independence in the context of conditional probability, the students may not have been able to correctly reason about the relationship between independence and mutual exclusivity without such a context. Perhaps this is what Kelly and Zwiers are arguing when they say that, "we, as instructors, do not take the time to relate the two concepts," referring to mutual exclusivity and independence (1988, p. 100).

Since independence can be defined outside of a conditional probability setting, it is not limited to an order of experiments. This can be seen in a standard deck of cards. The event of drawing a spade is independent of the event of drawing an ace. We can see this since P(♠) = ¼, P(ace) = 1/13, and $P(ace \cap ♠) = 1/52 = 1/4 \times 1/13$. This example demonstrates that independence can be accessibly thought of outside the context of conditional probability. This can prove useful since mutual exclusivity is also defined without respect to time. For instance, in the above case, it can quickly be demonstrated that $P(♥ \cap ♠) = 0$ since this intersection is empty; so these two events are mutually exclusive.

Altogether, we see that different researchers emphasize two different contexts for the independence of two events. Some focus on conditional probability for the definition of independence (Definition 1), while others focus on the product definition (Definition 2). Shaughnessey (1992) provides an explanation of why educators and some researchers focus on conditional probability when dealing with independence in that it is more intuitive for students to explore independence in the context of conditional probability. Regardless, students' misconceptions about the relationship between independence and mutual exclusivity prevail.

## Methods

### Overview

In order to explore my research question, I conducted a series of interviews with undergraduate students. Students from an introduction to proof course were solicited for participation because these participants would have higher-level (yet, still developing) proof schemes. The students were randomly selected from a group of volunteers, removing researcher bias from participant selection. I designed the series of interviews to ascertain a better understanding of how the students' proof schemes relate to their understanding of certain definitions. In order to investigate this relationship, the researcher focused on the four following areas:

1) Assessing the students' general proof scheme,
2) Investigating the students' understanding of mutual exclusivity and independence,
3) Students' assertion of a relationship between mutual exclusivity and independence, and
4) Students' attempt to prove the relationship they believe exists between mutual exclusivity and independence.

Each of these areas serves towards the research goal of improving mathematics education researchers' knowledge about students' understanding and misconceptions about mathematical definitions and how this relates to their general proof schemes.

Four students were randomly chosen out of volunteers to participate in the study, one of whom withdrew from the study for personal reasons. Three interviews were conducted with each of the three remaining participants (Alex, Betty, and Caroline) for a total of nine interviews. Each series of interviews focused on specific aspects of the research objectives. The first interview assessed each participant's proof scheme through his or her criticism of given proofs (see area 1, above). In the second interview, the participants discussed their understanding of the meanings of mutual exclusivity and independence and explored various sample spaces (see area 2, above). In the third interview, the participants asserted their understanding of the relationship between mutual exclusivity and independence and proceeded to prove this relationship (see areas 3& 4).

Each interview was audio and video recorded with multiple devices. All participant work was retained for further reference. After each interview, the researcher carefully watched the video recording and listened to the audio recording in order to gain insight into how the participants thought about the mathematics under discussion. The review sessions were used to generate models of the participants' thinking, gather quotes from the participants, and develop tasks that would allow a better understanding of the participants' thinking.

**Assessing General Proof Schemes**

We can distinguish between two general methods for assessing students' proof scheme in the available literature. Some researchers require that participants generate their own proofs in order for the researchers to evaluate the students' proofs schemes (e.g. Harel & Sowder, 1998; Housman & Porter, 2003). Other researchers focus on participants' verification or rejection of given proofs (e.g. Alcock & Weber, 2005; Martin & Harel, 1989). Research for the purposes of this study implemented the latter method. While this method allows sufficient opportunity to evaluate a participant's proof scheme, it can be argued that requiring participants to generate proofs could provide deeper, more meaningful insight into the participants' thought processes. Since this research aimed to obtain a more general (hence, less intensive) assessment of their proof scheme, participants were required to accept or refute presented proofs and then explicitly justify their reasoning. While this allowed insight into how the participants thought about proof, it did not allow the researcher to observe their own novel approaches to proving the mathematical relationships addressed.

The researcher presented each participant with a matrix of proofs (Appendix A). This matrix consisted of three different proofs (3 columns) for each of three different mathematical assertions (3 rows), a total of nine proofs. For each mathematical statement (row), the matrix contained one proof from each of the three levels of Harel and Sowder's (1998) proof schemes (External, Empirical, Analytical). The order of these proof schemes within each row was random so that no pattern could be inferred from previous rows. The matrix contained six proofs (rows 1 and 2) identical (or similar to) the proofs Martin and Harel (1989) used to assess preservice teachers' proof schemes. For some of the proofs, Martin and Harel (1989) did not provide examples of all proof schemes so the researcher created some of these proofs in order to reflect specific proof schemes. In collaboration with my advisor, I developed the other three proofs (row 3).

Each proof in row 1 claims to prove that "if the sum of the digits of a whole number is divisible by three, then the number is divisible by three." The participants were asked if each of the proofs convinced them of the relationship's truth. The first proof in this row uses an Inductive Empirical proof scheme by showing one example of a random large number that exhibits the desired relationship. It is important to note that, in this case, "Inductive proof scheme" does not mean the same as the principle of mathematical induction. Harel and Sowder (1998; 2007) use this term to denote students who believe that a finite number of random examples can imply the truth of a mathematical relationship for infinitely many examples.

The second proof in this row uses a Deductive proof scheme, where a given whole number "*d*" is written in expanded notation base 10 and rewritten in sigma notation. This summation is then separated into two summands so that one summand has a factor of nine and the other is the sum of the digits of *d*. This proof is relatively index intensive and quickly changed representations of *d*, causing it to be the most challenging proof of the entire matrix to verify as will be discussed in the next chapter. The last proof in row 1 exhibits an External proof scheme, more precisely the Ritual proof scheme. It applies mathematical induction so that it might seem that the result follows, but neglects that the sum of the digits of integers does not necessarily increase as integers increase (e.g. the sum of the digits of 21 is not three more than the sum of the digits of 18).

Each proof in the second row of the matrix asserts that "if a divides b and b divides c, then a divides c." The first proof in this row implements a Deductive proof scheme and applies the definition of divisibility in order to rewrite *b* and *c* in terms of *a*. Notice that none of the three elements in the proof are defined to be members of any set, and that the most general definition of divisibility is used. This allowed the interviewer to observe whether the participants had objections or reservations about the proof simply from the fact that the elements could be in any set. The second proof in this row is an Empirical proof that, as in the first row, uses a random selection of numbers and confirms that the relationship holds for that case, finally implying that this relationship generalizes. The third proof in this row uses a combination of External and Empirical proof schemes in that it addresses only two specific cases (Empirical) affirming the inverse of the argument (External) and showing the original argument. Notice that affirming the inverse of a statement does not affirm the original statement, much less for a specific case. This proof was used to determine whether the participants would assume the validity of proof based on logical argument (inverse, converse, contradiction, etc.) despite the logic's form and whether this validated the original statement.

In the third row in the matrix of proofs, each entry claims to prove that the sum of a triangle's interior angles is a straight angle. The first proof is a proof by picture that cuts the vertices from a given acute triangle and realigns them to form a straight edge. This is an Empirical proof. The second proof in this row uses the fact that two copies of the same triangle make a quadrilateral that, having interior angles that sum to 360°, has twice the interior angle measure of the original triangles. In this proof, the assertion of that the quadrilateral has interior angles that sum to 360° is never verified but rather is assumed to be true. This causes the proof to fall in the domain of External conviction since the rule cited has only the authority of the person who wrote the proof.

Similarly, the third proof in the third row has some ambiguity - although, one could argue, very little. This proof uses a line parallel to one side of a given triangle and passing through that side's opposite vertex. This forms two transversal lines of a pair of parallel lines. The congruency of alternate interior angles is then applied to show that the sum of the triangle's interior angles is 180° (the straight angle of the constructed line). While this proof assumes the congruency of alternate interior angles, this fact can be quickly verified. The proof of a quadrilateral's interior angle measure is equally labor intensive as the proof of a triangle's interior angle measure. Because of this, the third proof conveys a weak Analytical proof scheme.

**Investigating Participants' Understandings of Mutual Exclusivity and Independence**

The researcher conducted a second interview with each participant in order to understand how the participant thought about the concepts of mutual exclusivity and independence and about probability in general. In this interview, the participants began by giving definitions for and providing examples of mutual exclusivity and independence. The participants were then presented two sample spaces (Appendix B) and various events in those sample spaces. Within these sample spaces, the students were asked to determine whether pairs of events were mutually exclusive or independent and to explain their reasoning about each pair of events. After the two initial sample spaces were discussed, the researcher presented various improvised examples of sample spaces in order to further understand how the participants thought about the concepts of mutual exclusivity and independence and how the two are related.

The first sample space consisted of a standard deck of 52 cards and a fair six-sided die. The researcher explained to the participant that an event in this sample space consisted of drawing one card from the deck and rolling the die one time. The researcher initially presented two events: Event A- drawing a spade and rolling any number and Event B- drawing any card and rolling a three. After discussing whether events A and B are mutually exclusive or independent, the participants were then asked to consider a third and fourth event: Event C- drawing a spade and rolling a three and Event D- drawing a heart and rolling a four. We see that Events A and B are independent since $P(A)*P(B) = (1/4)*(1/6) = P(A\cap B)$. Also, Events A and B are not mutually exclusive, since their intersection is nonempty (Event C). We see also that Events C and D are not independent since $P(C)*P(D) = (1/24)*(1/24) \neq 0 = P(C\cap D)$. Furthermore, this last equality implies C and D are mutually exclusive.

Notice that this sample space consists of two physically separate sample spaces; we will call these “subsample spaces”. This sample space was chosen so that a single event could have two independent qualities that were not physically connected, as this could possibly have some influence over the participants’ responses. Admittedly, this seems to be a rather contrived scenario. But, the widely used (likely familiar) subsample spaces were chosen to allow the participants to more easily express their thoughts on how drawing a card and/or rolling a die occur and quickly calculate theoretical probabilities in their own verification of relationships in the sample space. Furthermore, in this sample space the only trait on the cards that the researcher presented was that card's suit. Because of this, the deck of cards was reduced to four traits (suit), each of which was on thirteen of the 52 cards. This contributed an interesting quality to this sample space that was unpredicted and that we discuss in the next chapter.

The next sample space consisted of a standard fair coin and a spinner. The spinner contained the colors blue, red, orange, green, and yellow that occupied 30, 20, 20, 20, and 10 percent of the spinner's area, respectively. The spinner was designed to have outcomes with theoretical probabilities that would not be computationally taxing, but would still provide varying probabilities between events. The participants were then shown a list of events in this sample space (A- blue/heads, B- red/tails, C- green/heads, D- blue/tails, E- {yellow or blue}/heads, F- purple/heads) and asked to determine whether pairs of these events were mutually exclusive or independent. We notice that Events A through D are pair-wise mutually exclusive and pair-wise *not* independent, Event F is *both* mutually exclusive *and* independent with all other events in the list, and Events A and E are *neither* mutually exclusive *nor* independent.

In this sample space, participants could consider an event that was a subset of another event and an event with zero probability. Each of these scenarios is a special case of the relationship between mutual exclusivity and independence. This sample space is similar to the first sample space because it contains two subsample spaces. This sample space differs from the first in that both subsample spaces in the second sample space have the quality that performing an action does not “remove” that outcome from the sample space whereas drawing a card from the deck of cards in the first sample space physically removes that card from the deck.

Once the participant discussed several pairs of events in the second sample space, the researcher asked probing questions intended to provide a deeper understanding of the participants’ reasoning. The researcher also questioned about other, improvised sample spaces in order to explore special cases of mutual exclusivity and independence. For example, if a participant distinguished between replacing the card in to the deck of cards and keeping the card out of the deck of cards, the researcher posed more questions about events related to a deck of cards without the die involved. Another example might be to have the participant consider the events of rolling an odd number and rolling a prime number on a fair 6-sided die. This event was used in order to discuss two events that were neither independent nor mutually exclusive.

**Proving a Relationship Between Mutual Exclusivity and Independence**

In the third interview, the participants were asked to assert a relationship between the concepts of mutual exclusivity and independence. The researcher used the questions from Manage and Scariano (2010). Each question assumed that two events, A and B, were in the same sample space and had nonzero probabilities. The first question stated that A and B were mutually exclusive and gave four answer choices: A) “then A and B are independent”, B)“then A and B are not independent”, C) “then A and B may or may not be independent”, D)“this is not enough information to determine whether A and B are independent”. Once the participant responded to this question, the researcher gave him or her a copy of his or her responses from the second interview and asked the participant to provide a proof for the relationship and to explain his or her reasoning throughout the proof.

The second question stated that A and B were independent and also gave four answer choices: A) “then A and B are mutually exclusive”, B)“then A and B are not mutually exclusive”, C) “then A and B may or may not be mutually exclusive”, D)“this is not enough information to determine whether A and B are mutually exclusive”. Again, once the participant responded to this question, the researcher asked the student to provide a proof for the relationship and to explain their reasoning throughout the proof.

Once these proofs were completed, the researcher probed the participant’s understanding of mutual exclusivity and independence based on the assertions that the participants made. This was mostly so the researcher could allow the participant to elaborate on their thought process, but this time was also used to provide examples of events in sample spaces that would perturb the participant’s conception of the two ideas. These tasks were developed based on analysis of the previous interview and some tasks were explored in response to participant examples.

**Data Analysis**

The researcher analyzed each video and audio recording after each interview in order to determine participants’ proof schemes, identify major themes of thinking, model participants’ understanding of mutual exclusivity and independence, draw quotes from the dialogue, and develop individualized tasks that would help explore each participant’s understanding in the subsequent interview. The researcher compiled various video clips that supported or challenged the researcher’s model of the participants’ understanding. These clips were then viewed several times in order to design tasks that would allow the researcher to resolve discrepancies between his models and the participants’ actions. The researcher relied heavily on Harel and Sowder’s (1998) and Martin and Harel’s (1989) examples of assessing proof schemes after the first interview with each participant.

Analysis for each interview required repeated video observations. In initial viewings, the researcher documented general responses to questions and tasks and assessed the responses' mathematical correctness. In the second viewing, the researcher focused on the participants' explanations and reasoning for each response. The researcher then compared responses and reasoning between tasks in order to develop holistic models of participants' thinking. Overarching motifs and themes between responses were supported by phrases and reasoning that the participants repeatedly used or referenced as well as quotes that exemplify and illustrate themes.

## Analysis

### Alex

Alex, the first participant in the study, was enrolled in an Introduction to Proofs course. At the time, he had not taken any Probability or Statistics course at the undergraduate level. A mathematics major with interests in physics and engineering, Alex also seemed very interested in mathematics education research and said that he felt that this would be an interesting introduction to how the process works. Here is an anecdote to convey Alex's personality: when asked to provide a pseudonym for the research project's online schedule, he used the name "Grigori Perelman."

#### Interview 1

Throughout the interview, Alex exhibited a predominately Analytic proof scheme. Eventually he correctly supported all Deductive proofs and refuted all Empirical and External proofs, citing appropriate flaws in logic or reasoning. In a few instances, he showed signs of relying on a proof's form rather than content, signifying an occasional tendency toward a Ritual (External) proof scheme. Alex was also very pedantic about precise details, reflecting a skeptical point of view and checking for logical progression in each proof.

Alex displayed a deep understanding of examples and their use in proof. This was quickly evident in the first example in the matrix of proofs. This proof, applying an inductive (Empirical) proof scheme, used an example of a large random number that exhibited the desired result. After reading the argument, Alex immediately said, "Yeah, this is bogus." He later refuted the second proof of the second row, another inductive proof very similar to the first. Similarly, Alex rejected the first proof in the last row, a proof by picture, in which angles from one triangle were removed and realigned to show a sum of 180°. He stated that the picture seemed to show the relationship in this triangle, but that the picture failed to generalize the relationship to all triangles. These examples highlight Alex's ability to refute proofs that inappropriately use examples.

Alex also remained skeptical of the motivations within each proof throughout the interview, which was filled with long pauses of two to three minutes during which Alex contemplated a single line of an argument. This trait of his proof scheme was most evident in his examination of both the second and third proofs on the first row. Interestingly, these were the only proofs in the matrix about which Alex changed his mind regarding their validity. Combined, Alex spent more than half of the one-hour and three-minute interview on these two proofs (fifteen minutes on the second and eighteen minutes on the third).

His time on these proofs was spent mainly evaluating the justification from each row to the next and explaining his reasoning to the interviewer. Throughout his discussion of the second proof, he questioned the author's motivation for changing notation from $10^k$ to $9*p_k + 1$ and then later questioned the phrasing of the final argument even though he had articulated the general idea of the proof immediately before that. Alex explained that, because of this skepticism, he initially believed the proof was invalid before reading through the proof again and declaring it sufficient.

Alex also used examples to verify and refute symbolic relationships. For example, in the second proof on the first row, Alex used the index 2 in order to better visualize the process of expanding $10^2$ using the notation presented ($100 = 9*p_2 + 1 = 9*11 + 1$). Although this did not show the relationship for every case of $10^n$, Alex saw the pattern in the process. His value of counterexamples was evident in the third proof of the first row when he refuted part of the proof's argument by showing $[\Sigma(\text{digits of } 9)]+3 = 12 \neq 3 = [\Sigma(\text{digits of } 12)] = [\Sigma(\text{digits of } 9+3)]$.

It should be noted however, that Alex initially supported this proof, although weakly as he said he was "going back and forth between believing it and not believing it." It was not until the interviewer asked for an explanation for this part of the argument that Alex attempted to find examples of $[\Sigma(\text{digits of } 3k)]+3 \neq [\Sigma(\text{digits of } 3k+3)]$ for $k \in \mathbb{Z}$. This also reflects Alex's focus on the form of a proof, which is indicative of a Ritual (External) proof scheme especially since, after seeing the first proof of the relationship, Alex had said he would probably use mathematical induction to prove this relationship. This helps explain why he initially accepted this proof, since he might have been expecting a proof by induction to be valid.

Another interesting aspect of Alex's proof scheme is his emphasis on the axioms of real numbers and considering the space in which he was working. These qualities were evident in the second proof on the first row, the first proof on the second row, and the third proof in the third row. In the first two cases, Alex explicitly applied closure of integers under addition and multiplication. In the second case, Alex also invoked the associativity axiom for real numbers. In the third case, Alex suggested that the sum of the interior angles of a triangle might not be 180° in non-Euclidean space. While this could be a manifestation of the rigor required in his Introduction to Proofs course, it is evident from these examples that Alex had internalized a mindset that considers the system in which a proof is argued and its fundamental axioms. It should be noted that Alex's use of axioms in this interview reflects Harel and Sowder's (1998) Intuitive-Axiomatic proof scheme.

These qualities of Alex's proof scheme combine to support his overall emphasis on the form of a proof. This emphasis is manifested in his tendency toward induction, pedantic and specific use of axioms, and refutation of proofs by example. This attitude was also reflected when Alex said, "Graphical proofs in general can be misleading. So I generally just wouldn't trust them from the start." We see Alex uses the form of a proof to make initial judgments, but his skepticism forces him to evaluate a proof based on its line-by-line merit. From this, we can conclude that Alex generally displays an Intuitive-Axiomatic (Analytical) proof scheme with tendencies toward a Ritual (External) proof scheme.

**Interview 2**

Alex explored two sample spaces and discussed a few other examples that he used to help describe his understanding of mutual exclusivity and independence. As we will see, Alex displayed an extremely internalized and powerful conception of independence. Alex defined independence as, "[when] the outcome of one event does not affect the outcome of a subsequent event." This definition implies an emphasis on a sequence of events, where one of the events being considered must occur prior to the other.

With regard to mutual exclusivity, however, Alex was less certain of a formal definition — changing his definition twice throughout the interview until eventually declaring, "Performing an event or series of events causes a subsequent event to have zero probability of happening." Again, Alex implies that this relationship is defined over a period of time. It is important to note that Alex's initial definition of mutual exclusivity (consistent with the mathematical definition) was not defined over time, but rather instantaneously. It was not until he had considered examples in the two given sample spaces that he changed this definition to more closely resemble his definition of independence.

When prompted for an example of independent events, Alex gave two examples: a die and a coin. He stated that rolling a six on the first roll of a die does not affect rolling a six on the second roll of a die and gave an analogous explanation for the coin. These examples are consistent with his definition of independence, implying that the two events in consideration take place at separate times. His initial examples of mutually exclusive events exhibited what he described as “well-defined states” including raining versus not raining, sides of a die (“you can’t roll both a 5 and a 1”), and a coin (“it’s either heads or tails”). These examples support his original definition, which considers the two outcomes instantaneously in that it cannot both rain and not rain at the same time.

Later in the interview, after changing his definition of mutual exclusivity, Alex used the events given in the first sample space to make a new example, describing the process of repeatedly drawing “any card” without replacement until all spades were exhausted. In this case, drawing “any card” and drawing a spade were mutually exclusive since drawing “any card” can eventually cause drawing a spade to have probability zero. This example seems much more convoluted than the first three examples, but supported Alex’s newer definition of mutual exclusivity. We can see that Alex’s conception of independence was so strong that it not only influenced how he defined mutual exclusivity, but also caused him to reject three different examples and develop a new concept image for mutual exclusivity wherein one event must cause a subsequent event to be impossible. This new concept image was so strong that, when asked to reconcile this new definition with his original examples, Alex reneged on their mutual exclusivity (i.e. heads on a coin does not cause “not tails” later).

Equally intriguing is the fact that Alex independently asserted a corollary to his new definition of mutual exclusivity. In this corollary, Alex stated that if the two events are mutually exclusive, then they cannot be independent. This reflects the (almost) exact relationship outlined in Manage and Scariano (2010) and investigated in the third interview of this thesis research. Alex used an explanation analogous to that described in Manage and Scariano (2010). He asserted that, since one event causes the second event to have zero probability, the first event changes the probability of the second event and therefore the two events are not independent. It should be noted, however, that Alex did not consider the case when the second event in the sequence already had zero probability until later in the interview.

In the first sample space (a standard deck of cards and fair die), when Alex was asked to compare events with respect to mutual exclusivity and independence, he focused on whether the first event replaced the card into the deck. In each comparison, this subsample space seemed to have a greater effect on the independence or mutual exclusivity of the two events. When asked about this, Alex responded, “the die… it really doesn’t matter what you do because, uh it’s just a property of rolling a die you… when you... you can do it as many times as you want and every sing… If you’re just rolling dice over and over, those are always independent.” Because of this focus on the deck of cards, any pair of events could be both independent and not independent — the only difference being whether the card was replaced after the first event. Alex changed his definition the first time he compared events in this sample space with respect to mutual exclusivity. Because of this, his focus still remained on the deck of cards.

In the second sample space (a spinner and fair coin), Alex found every event to be independent and “not mutually exclusive” of the other, since the occurrence of one event could not change the probability of a later event, much less cause it to be zero. In every case, he contended that no replacement needed to be considered since nothing was removed from either the spinner or the coin. This reflects his explanation of the die in the first sample space.

One pair of events in this sample space caused Alex to slightly modify his definition of mutual exclusivity. Alex was asked to consider event A (spinning blue and tossing heads) and event F (spinning purple and tossing heads). This pair of events was different from the previous pairs in that purple was not on the spinner and so event F had probability zero. Alex had asserted that these two events were independent (as were all other events in this space), since event A could not affect event F. But, according to his corollary, Alex knew that if two events were mutually exclusive, then they were not independent. He argued that these two events cannot be mutually exclusive since they were independent so he emphasized that event A did not *cause* event F to have probability zero. This allowed him to reconcile the pattern that every event in this sample space was both independent and mutually exclusive. This also provides another example where Alex's definition of independence changed his definition of mutually exclusive events.

In this interview we have seen that Alex's concept definition of independence relies on a temporal relationship in the sample space under consideration. We see that this is also reflected in his concept image of independence, as exemplified by his discussion of the sample spaces. If the set of objects considered in the sample space are "removed" with the occurrence of an event, then replacement of the object implies independence. Otherwise, the events are not independent. This conception of independence is so powerful in Alex's discussion of probability that his definition of mutual exclusivity changed to accommodate this with/without replacement system.

**Interview 3**

In this interview, Alex was asked the questions from Manage and Scariano (2010). Recall the two multiple-choice questions: each supposed that two events are mutually exclusive and independent, respectively. In the first question, the respondents are given four answer choices: A) A and B are independent, B) A and B are not independent, C) A and B may or may not be independent, and D) there is not enough information to determine whether A and B are independent. In the second question, similar choices are given, the exception being that "independent" is replaced with "mutually exclusive."

In Manage and Scariano (2010), the events are also given to have nonzero probabilities. In this interview, the questions were initially posed without the information that A and B had nonzero probabilities. This was because of Alex's explanation in the second interview when one of the two events had probability zero. In the first interview, Alex deliberately and repeatedly focused on the need to prove relationships for all cases. This question was posed initially to examine whether Alex would think to consider "all possible situations" which emphasized so much in the first interview. The interviewer then asked Alex to answer the question assuming that neither event had zero probability.

When asked the first form of the question, Alex asserted that mutual exclusive implied not independent. This claim was made using his final definition of mutual exclusivity. He directly referenced his own corollary from the second interview in which he made this exact assertion. When asked about the relationship when neither event had zero probability, Alex changed his response to answer choice D, "there is not enough information to determine whether these events are independent." The interviewer pointed out that Alex had definitely concluded an inverse relationship between the two terms, that information was added, and that this answer choice was less specific than the first response. Alex justified this change in his answer choice by explaining that under his definitions everything is highly dependent on context, referring to the sample spaces in the second interview. Although this explanation did not explicitly address the addition of more information to the problem, this does provide insight into Alex's own understanding that his definitions were useful only if the sample space containing A and B were known.

The interviewer then asked Alex the second question, keeping the information that A and B had nonzero probabilities. Alex responded with answer choice B, that the two events were not mutually exclusive. He supports his answer choice by saying, “one event’s not affecting the other event at all so, I mean, it’s not going to cause it to have zero probability cause it’s not changing the probability of the next event.” As with the first question, this answer choice supports the relationship between the mathematical definitions of independence and mutual exclusivity, despite Alex’s concept definitions that rely on the occurrence of A chronologically before B.

#### Comparing Alex’s Responses Across Interviews

From these responses, we see that Alex’s proof scheme played a lesser role in his justification of the relationships between mutual exclusivity and independence than his definitions. This is evident in his assertions of the relationship with and without the zero probability cases, which reflect a diminished emphasis on proof for every case when compared to his responses in the first interview. Since the zero probability case is equivalent to two events being both independent and mutually exclusive, this case is very important in the assertion of any relationship between the two concepts.

The results from these interviews also show that one can assert a correct relationship between two mathematical concepts despite incomplete or incorrect definitions. Alex’s corollary was a nearly logical proof that mutual exclusivity precludes independence, the “nearly” in reference to the zero probability case. This reflects the idea claimed by Kelly and Zwiers (1988) — that zero probabilities are “pathological cases” — since Alex did not consider these cases despite his previous emphasis on proof for all cases.

### Betty

Betty, the second participant in the study, was enrolled in the Introduction to Proofs course and in an undergraduate statistics course. An engineering major, Betty volunteered for participation in the project because she said she was interested in the idea of contributing to mathematics education research. Betty seemed somewhat tentative in her responses throughout the interviews, often questioning the interviewer whether her responses were correct. This could have been an artifact of her personality rather than her confidence in the mathematical process since Betty seemed more confident when she explicitly explained relationships between mathematical concepts, a phenomenon that supports the idea of proof as a sense-making process embodied by Harel and Sowder's (1998) process of ascertaining.

**Interview 1**

Betty displayed a predominately Analytic proof scheme with the exception that she accepted the third proof (External-Ritual proof scheme) in the first row of the matrix of proofs and accepted the third proof in the third row based on its mathematical appearance. Betty correctly refuted the three examples of Inductive (Empirical) proofs, but accepted one deductive proof because it "seem[ed] more mathematical." Her refutation of the inductive proofs shows her understanding of the importance of a general proof for all cases. Betty's acceptance of a proof based on it's seeming mathematical qualities and acceptance of the false proof by the principle of mathematical induction, however, indicate a tendency toward External (Ritual) and Empirical (Perceptual) proof schemes.

Betty also pointed out an inconsistency in one of the indices in the first row's second proof. This shows an insistence on understanding very specific aspects of a proof rather than drawing any assumptions about the proof's process. We can also see this reflected in her discussion of this proof when she said, "It's convincing now that I've worked it out, but as we saw I had problems, like, getting… following their path. So, I mean if [this person] could like talk me through it a little better, then I feel like it would have been a little bit easier to understand." Betty also contends that an explanation of connections between statements could help her better understand the proof even though she eventually accepted the proof as valid.

In the third proof on the first row, however, Betty was more lenient on the argument. She quickly accepted the proof by mathematical induction. Here, she may have been more preoccupied with the form (or "look") of the proof, rather than its mathematical validity. This idea was supported when Betty stated that her class had recently discussed the principle of mathematical induction. When asked which of the three methods she preferred, Betty chose the last proof because, she said, the base-10 expansion in the second proof was not obvious to her at all. This reflects a need to understand connections in a proof, even though this need was temporarily suspended in the case of mathematical induction. This need was also addressed later in the interview, when Betty described the process of verifying for herself relationships she felt she did not understand in class.

Betty accepted the first proof in row 2, which used an Analytical (Axiomatic) proof scheme to prove "a divides b and b divides c implies a divides c." Her only concern with this proof was the use of the word "divisible" which she thought implied that the numbers under question must be integers. Betty refuted the other two proofs in this row, which exhibited Inductive and Ritual/Inductive proof schemes, respectively. She rejected these proofs with little hesitance. When asked whether the third proof could be giving a counterexample, Betty said, "There *are* more possibilities," referring to negating the original statement. She went on to describe other combinations of negating the hypotheses of the statement and also compared this proof to the second (Inductive) proof. This shows that Betty has a good understanding of negating statements, but does not show how much or if Betty values counterexample.

In the third row of proofs, Betty quickly refuted the first two proofs. She claimed that the first proof only showed one case and did not completely prove the relationship since the constructed angle could not be accurately measured. This reflects her ability to identify false proof by example. In the second proof, Betty questioned the ability to apply the fact that a quadrilateral has interior angles that sum to 360°. This highlights Betty's ability to question what inferences can be made without justification, citing that the results of the proof in question are used to prove the feature of the quadrilateral. Betty's reasoning in the third proof, however, reflects an emphasis on the ritualistic aspect of the proof, rather than the content of the proof. These examples show a healthy skepticism of Authoritarian proof and also a tendency toward Ritual proof, both of which are External proof schemes. Additionally, Betty's explanations in refuting Inductive proof schemes support an emphasis on proof for all cases.

**Interview 2**

When asked what it meant for two events in a sample space to be independent, Betty responded, "The intersection is zero. Is it? That's what I'm asking. I don't remember."
Betty almost instantly changed this to, "Two events are independent if the probability of A occurring does not affect the probability of B occurring." Betty then described the independence of events A and B using the equation $P(A) = P(A|B)$. Neither of these representations necessarily implies a chronological distinction between events A and B (as was seen with Alex's use of "subsequent"). But, when prompted for an example of independent events, Betty described the act of picking a card from a deck of fifty-two cards and putting it back so that the probability of picking a second card is not affected. Similarly, when asked for an example of events *not* being independent, Betty provided the case of picking a card and not replacing it. These examples are

consistent with a conception of independence in the context of a “with replacement” and “without replacement” conditioning event.

In contrast, Betty defined mutually exclusive events with the statement, “you can’t have both at the same time.” This definition explicitly states that the events can be compared instantaneously. Here, Betty gave the example that the queen of hearts and jack of diamonds are mutually exclusive, since they cannot both occur when one card is drawn. We notice that this definition is consistent with the mathematical definition and that this example is consistent with Betty’s definition. Betty did spend much more time defining mutually exclusive events compared to her definition of independence, but once she determined this definition, she held firm to its accuracy saying, “I’m sorted now.” This reflects her need in first interview to prove relationships in order to understand them.

Betty’s initial confusion of independent events as events that “don’t happen at the same time” reflects the most common misconception in Manage and Scariano (2010). Although she quickly changed her mind about the definition of independence, this confusion was apparent in her use of mathematical notation to represent the ideas (discussed below). Also, when explaining her conditional notation of independence, Betty described two independent events as “completely separate,” which one could argue is a descriptor more applicable to mutually exclusive events since their intersection is empty.

More than once, Betty wrote an equation involving probabilities and quickly erased it saying, “That’s just something I remember from probability.” For instance, she initially used “$P(A\cap B)=0$” to represent independence and used the equation “$P(A\cap B)=P(A)*P(B)$” to define mutual exclusivity. As mentioned, these equations were quickly erased. The former, however, was eventually used to describe mutual exclusivity. The latter, Betty admitted, “I have no idea where that came from or if that's even mutually exclusive. And I would not be able to come up with [it]." We notice that Betty's second description of independence, P(A)=P(A|B), is true unless the probability of B is zero. In this case, the statement P(A|B) makes no sense. One could adapt this statement to say, "two events A and B are independent if both have nonzero probability, P(A)=P(A|B), and P(B)=P(B|A)."

When discussing independence in the first sample space, Betty considered one event chronologically after the other and focused almost entirely on whether the card was replaced after each event, judging whether the first event affected the second. One interesting aspect of her methods during this discussion is that she also considered the order in which the events occurred. This was important when considering event B without replacement because drawing "any card" could remove a spade from the deck, affecting the probability of event A, and causing these two to be not independent. Conversely, if event A was successful first, event B could still occur since only one card was removed and "a card" could still be successfully drawn. This implied that order of events as well as replacement mattered in this sample space.

In the second sample space, Betty identified all given events as independent of each other. When asked whether she could come up with a pair of not independent events, she said that this was not possible since one spin won't affect a second spin and one toss of a coin won't affect another toss of the coin. This, combined with her emphasis on replacement in the first sample space, explicitly shows that Betty's conception of independence is entirely within the context of the sample space. In contrast, when considering mutual exclusivity in both sample spaces, Betty correctly identified events with empty and nonempty intersections as mutually exclusive and not mutually exclusive, respectively. This reflects the consistency between Betty's conception of mutual exclusivity and the mathematical definition.

In this interview, we see that Betty's concept definitions, though initially inconsistent, are each strongly internalized when evaluating the independence and mutual exclusivity of specific events in specific sample spaces. This is evident because, once Betty defined each term, she was "sorted" on how to verify them and seemed to develop quick checks in order to do this (i.e. "Can these happen at the same time?"). Her spoken reasoning for two events' independence and mutual exclusivity reflected these quick checks. One interesting aspect of this interview is that Betty seemed to have little problem with the same two events being both independent and not independent in the first sample space. One could argue that independence was seen as a consequence of actions on the sample space and not a mathematical trait of the two events themselves since both conditioning and order affected two events' independence.

Conversely, every event in the second sample space was independent of every other event. This was reflected in Betty's claim that no two independent events could be found in this sample space. We see from this that Betty's conceptions of independence and mutual exclusivity allow for at least two types of sample spaces: one in which replacement affects independence and another in which independence occurs between all pairs of events. This is a reflection of the occurrence of the two events at different times.

**Interview 3**

Recall that the two multiple-choice questions in the third interview asked the participants whether mutually exclusivity implied independence and whether independence implied mutual exclusivity, respectively (with additional choices logically equivalent to "maybe" and "not enough information"). When presenting these questions, the interviewer intentionally left the phrase "A and B have nonzero probabilities" out of the question statement. This is because, in the previous interview, Betty's response to events with zero probability did not differ from other events. The interviewer intended to ask each question with two parts, similarly to the third interview with Alex. As we see, Betty's responses to the questions precluded the need for adding this information.

In response to each of the two questions, Betty concluded that there was not enough information about the sample space and that two mutually exclusive events can be both independent and not independent — leading her to respond with both answers C and D. Betty answered almost identically to the second question. She explained that in the previous interview she had seen mutually exclusive events that were both independent and not independent (a copy of her responses from the second interview was presented to her). She also explained that she saw independent events that were both mutually exclusive and not mutually exclusive in the second sample space. From this, Betty reasoned that more information was needed about both the sample space and the actions taken between the occurrence of the first event and second (e.g. replacement, non-replacement). Again, we see independence is affected by the context in which the events take place.

**Comparing Betty's Responses Across Interviews**

Betty's proof scheme showed that she is more inclined to want to verify mathematical relationships on her own. This was evident as she "sorted" herself about the definitions of independence and mutual exclusivity. During this process, Betty successfully reconciled her definitions of the terms with symbolic representations (about which she was admittedly unsure) that she had recalled from her statistics course. Betty used these definitions to investigate the sample spaces in the second interview, the results of which had a direct affect on her reasoning in the third interview. Because Betty's definition of independence relied so heavily on the sample space and whether replacement occurred, she had examples of all different combinations of independence and mutual exclusivity.

After the third interview, Betty and the interviewer discussed mutual exclusivity and independence. The interviewer posed the question, "How do you find the probability of two things happening at the same time?" This prompted an hour-long discussion about independence. In this discussion, Betty independently produced the formula, "$P(A \cap B) = P(A)*P(A|B)$." After this, she logically validated her conditional notation of independence: $P(A) = P(A|B)$. Betty concluded that independence meant $P(A \cap B) = P(A)*P(B)$.

In exploring a few more sample spaces and examples of independence the discussion became focused on comparing one outcome in a sample space versus two outcomes. This distinction of considering two events when completing one physical task allowed Betty to adapt her conception of independence from a sequence of chronological events to a simultaneous comparison. The interviewer used guided questioning to ensure that Betty discussed the equations $P(A \cap B) = 0$ and $P(A \cap B) = P(A)*P(B)$ in quick succession. From this, Betty concluded that independent events are mutually exclusive only if one or both of the events has zero probability.

**Caroline**

The third participant in the study, Caroline, was taking the same Introductory to Proofs course as the other participants. At the time of the study, she had not taken an undergraduate statistics or probability course. Caroline was majoring in mathematics education and volunteered for the study so that she could get free tutoring in a Differential Equations course. Caroline exhibited a combination of all three proof schemes — External (Authoritarian), Empirical

(Inductive), and Analytical (Contextual) — at various points in the interviews. She occasionally emphasized form over content and once relied on a proof's conclusion in the proving process.

Caroline displayed a conception of independence that varied according to context although she was confident that independence meant that two events "don't affect each other." She also had initial difficulty defining mutual exclusivity, eventually settling on a definition logically equivalent to the formal definition, which she later used to accurately identify pairs of mutually exclusive events. After exploring extra sample spaces and considering different contexts of events, Caroline concluded one direction of the relationship between mutual exclusivity and independence — that two mutually exclusive events cannot be independent if the events are compared within one "action." Caroline did not, however, consider the zero probability case acknowledged in the literature.

**Interview 1**

Caroline refuted three proofs outright. Two of these three proofs — the first proof in the first row and the second proof in the second row — exemplified an Inductive (Empirical) proof scheme. Caroline argued that the first two proofs are incorrect because a general statement is not proven by one example. She explained, "It's correct, but even though I know that's a rule, this wouldn't convince me," and later said, "You can't just use one example." The third proof that Caroline refuted based on argumentation was the third proof in the second row, which used a combination of an Inductive proof scheme and a Ritual proof scheme. Caroline recognized that this proof did not logically conclude the intended result, but did recommend that the prover might attempt a proof by contraposition or another logical method and noted that this proof was not by contraposition. She then attempted to prove the relationship by contraposition, but used the original statement to prove the contrapositive. This shows the first of three specific examples where Caroline uses the statement she wishes to prove in order to justify a proof.

The next two examples where Caroline uses the results in her argument for the validity of the proof are the first and second proofs in the third row. In discussing the first proof, where a triangle is shown to have interior angles summing to 180°, Caroline says, "I would believe that… Just because from a… Just thinking about it and just knowing a triangle and *knowing* that the interior angle, the sum of them *will* be 180 degrees." She goes on to say that "this proof itself, doesn't… I mean it doesn't give you an opportunity to prove that it wouldn't work — to disprove it." In the second proof on this row, which builds a quadrilateral from two copies of the same triangle, Caroline again invokes the fact that she "knows just from basic geometry" that quadrilateral has 360 degrees and a triangle has 180 degrees to support her reasoning that the second proof is valid. These last two examples reflect a combination of Authoritarian (External) and Contextual (Analytic) proof schemes because she fails to remove herself from the context of her current understanding of mathematics while relying on the authority of "facts" she learned in "basic geometry."

Caroline also shows a Ritual (External) proof scheme throughout the interview. The first instance of this occurs when Caroline notes the lack of a formula in the third proof in the first row. This proof uses the Principle of Mathematical Induction, which Caroline says, "usually has a formula." Regardless of this discrepancy, Caroline still accepts the proof as valid. Another example of this is evident in the first proof in the third row with Caroline's emphasis on the "look" of the proof when she says, "that doesn't prove it with numbers, that just proves it… *literally* proves it." Later, in the third proof on this row, Caroline claims that the proof seems more formal than the first since it uses opposite interior angles.

Caroline found the computations too difficult in the second proof in the first row to either accept or refute it. This proof had also caused problems with the other participants in the study and was by far the most computational proof in the matrix. This shows that Caroline is skeptical of the prover's abilities in some cases. Although the proof is logical, neither its form nor the seeming authoritarianism of the taxing computations persuaded Caroline. This contrasts with her tendency toward an Authoritarian proof scheme in other instances. One possible explanation for this discrepancy could be that Caroline was more familiar with the form of the proofs (i.e. Principle of Mathematical Induction) that she accepted as valid.

Caroline did accept the first proof in the second row (an Analytic proof scheme), declaring, "This is the exact way I would have done it." This was one of the two Analytic proofs that she accepted, the other being the third proof in the third row, which it seems appealed more so to Caroline via a Ritual proof scheme. However, this shows that, with at least some of the proofs, Caroline was able to distinguish more Analytical arguments from External and Empirical proofs.

We have seen that Caroline displays all three of Harel and Sowder's (1998) major levels in her proof scheme. Although Caroline accepted two of the three Analytic proofs, she also accepted one Empirical proof and two External proofs. In more than one instance, Caroline used the "look" of a proof as at least part of her justification in either accepting or refuting it. This is evidence that she showed tendencies toward Authoritarian and Ritual proof schemes. But her ability to identify and refute some, though not all Inductive proofs also showed a tendency to emphasize the need for proving a relationship for all cases. Furthermore, her rejection to either support or refute a proof for its complicated structure shows at least some tendency away from an authoritarian proof scheme.

**Interview 2**

Caroline's definition for independence was similar to the other two participants, stating, "Two events are independent if they do not affect each other." Caroline's examples of independent events were interesting in that Caroline described "everyday events" rather than "artificial" events (such as dice or cards) that are typically investigated in the classroom setting. For example, Caroline described how "the probability of someone wearing a red shirt is independent of their age." Similarly, when prompted for an example of events that are not independent, Caroline provided the example of someone who is forgetful is less likely to win a student lottery for a football ticket, since they are less likely to enter the lottery. This example implies a directly causal relationship, where the lower probability of the first event (entering the lottery) decreases the probability of a later event (being selected in said lottery).

When asked to define mutually exclusive events, Caroline seemed at a loss. Initially, she said, "they're separate from each other," in addition to admitting that she feels that the concept is similar to independence. After a short explanation of the words "mutual" and "exclusive," Caroline asserted that mutually exclusive things don't include any part of each other. As the interview progressed and the two main sample spaces were explored, Caroline seemed to become more comfortable with this second definition — progressively responding to questions more quickly and with greater confidence and giving more reasoning for her responses. These responses gradually became more consistent with the mathematical definition of mutual exclusivity.

When discussing independence in the sample spaces, Caroline questioned whether the "deck is reset" after each draw. As her exploration of pairs of events in this sample space continued, she repeatedly described the two cases of whether the cards are replaced in the deck.

In each case, resetting the deck yields independence and keeping the card out implies that the two events will not be independent. At one point, Caroline declares, "For the die, that one, it won't really have an effect because you're not taking anything away. You're just rolling it."

Caroline's reasoning in the second sample space was similar to her reasoning with the die in the first sample space. She identified mutually exclusive events accurately and quickly since she was able to determine whether their intersection was empty and compare that to her definition. She also found all events in this sample space to be independent. Her reasoning for this was that, "You're not getting rid of anything. Just because you land on blue doesn't mean you're taking away any of the blue." From this explanation, she again compares the two events in some chronological sequence.

**Interview 3**

In the third interview, Caroline initially thought that more information was required about the sample space in both problems (answer choice C). After comparing the events of rolling an odd number and an even number in the sample space of a die, Caroline considered two major cases: rolling the die once and rolling twice, "one after another." This is different from Caroline's consideration of with and without replacement in the first sample space of the second interview. Caroline explains that, "In this case [referring to sample space 2], you're saying like, two events. It's like spinning blue and flipping it on heads. Like, those are two obviously separate – like, happening at two different times – things. And this is getting – rolling even or odd – is two possibilities, two possible events that could happen by doing one thing."

Caroline then considered the two questions from Manage & Scariano in the context of these two cases. When considering the case of the two events being compared as one action, Caroline reasoned that mutually exclusive events would have an effect on each other. She did not, however, make a definite conclusion in the reverse direction nor did she acknowledge the zero probability case. The second case was not as simple, as Caroline explained, since they would be independent on a die but with a deck it would depend on replacement of the first event. When responding to the second question, Caroline discussed several scenarios until eventually deciding that more information was needed to answer the question.

#### Comparing Caroline's Responses Across Interviews

Caroline's proof schemes reflected all three of Harel and Sowder's (1998) major proof schemes. In assessing this, we saw that Caroline is sometimes persuaded to accept or refute an argument based on its form (or "look"), what she believes or knows to be true, and logical and deductive reasoning. These various levels of proof schemes were also evident in the arguments she gave to support her responses in the second and third interviews. The fact that Caroline did not consider the zero probability case in the third interview can be viewed as either an example of her tendency to occasionally accept an Inductive proof scheme (which does not consider all cases) or an artifact of her concept image for mutual exclusivity and independence.

### Comparing Alex, Betty, and Caroline

Alex, Betty, and Caroline each exhibited very different proof schemes. Alex's proof scheme was mostly Analytical (both Axiomatic and Transformational), with occasional emphasis on Ritual aspects of a proof. Betty's proof scheme was similar to Alex's — with the exception that Betty accepted an External (Ritual) proof and discussed proofs with less focus on axiomatic aspects but comparable focus on transformation. In contrast, Caroline exhibited all three major types of proof scheme. During the second and third interviews, each participant's proof scheme became more apparent as they explained their reasoning about mutual exclusivity and independence.

We saw that Alex’s conception of independence was so strong that it influenced his definition of mutual exclusivity. It can also be argued that Caroline’s definition of independence was affected by her conception of mutual exclusivity, since she simultaneously (or occurring with “with one draw”) compared events with respect to mutual exclusivity. In the second interview, Caroline judged independence by considering one event’s success following the other’s success. It was not until the third interview, when Caroline was trying to compare the two concepts that she then considered independence of two outcomes on a single action (i.e. a single card successfully satisfying A = ♠ and B = 8). It is not immediately clear why Caroline adapted independence to the single event instead of changing mutual exclusivity to sequential events as Alex did.

Each participant pointed out the importance of whether or not replacement occurred after drawing the “first card” from the deck of cards in the first sample space. All three participants also ignored the die in the first sample space when considering two events’ independence. Each of these phenomena can be attributed to a conception of independence that considers one event chronologically before the other event. Similarly, all three participants argued that no two events could be “not independent” in the second sample space from interview 2. From this conception of independence, the type of sample space and actions between each success of the two events can influence their independence. Caroline at least partially reconciled this conception of independence through her exploration and consideration of various sample spaces.

# Results and Conclusions

## Evidence Relating Proof and Definition

We see that proof schemes can be both restricted and enhanced by students' definitions of the mathematical ideas they consider. Though her reasoning was logically based on her previous experiences in the samples spaces, Betty's conception of independence and mutual exclusivity caused her to require more information about the sample spaces in question, in turn restricting her ability to draw conclusions between the two concepts. On the other hand, Alex and Caroline's ability to adapt their concept images and concept definitions allowed them to logically conclude one or both directions of the relationship between mutual exclusivity and independence, however correct or incorrect their definitions may have been.

In his proof, Alex claimed from his concept definition of mutual exclusivity that each mutually exclusive event would *cause* the other to have zero probability. This would make the two events "not independent" since his definition of independence necessitated each event to "not affect a subsequent event." Using similar reasoning, Alex concluded that independence implied "not mutual exclusivity." It should be noted however that, despite Alex's focus on "proof for every case" in the first interview, that he failed to assert a relationship for the case when one or both events were given to have zero probabilities. The contrast between his assertions about proof and his actions in proving this relationship reflects the "pathological" nature of zero probability cases pointed out by Kelly and Zwiers (1986). Interestingly, this also points to a characteristic of his definitions that may have influenced his thought process: an event with zero probability cannot "happen first" and therefore can neither *cause* nor *affect* any other event, as the definitions require.

Betty's concept images were so strong that logical reasoning resulted in her inability to assert any certain relationship between the two concepts. More specifically, Betty's personal experiences in the sample spaces allowed her to provide counterexamples to any explicit relationship between the two concepts. Since specific characteristics of sample spaces affected two events' independence, she required information about a sample space in order to make inferences about the events in question. This prevented Betty from generalizing an explicit relationship between mutual exclusivity to all cases, which her proof scheme required.

As mentioned, Caroline was able to adjust her conception of independence so that two events can be considered as results of a single action. This allowed her to conclude that mutually exclusive events cannot be independent after considering a few examples and reasoning through the relationships therein. This conclusion was based on a context and definition consistent mathematical definitions. Her proof of the relationship reflects her variety of proof schemes since she uses examples, indicating an approach from an inductive proof scheme, but then refers to her definition to generalize this outside of that specific context into an overarching logical conclusion. Caroline had difficulty asserting an implication in the second question using her case of comparing events in a single trial. In this instance, Caroline's lack of a coherent concept image in this newer context prevented her from concluding a relationship in this direction.

Recalling the Alex and Betty's general proof schemes (mostly Analytical and Analytical with Empirical and External tendencies, respectively), we consider how these related to their use of definition. Alex's dynamic concept image and unsolicited production of the lemma for the definition of mutual exclusivity reflect an Analytical frame of mind that is also geared toward finding asserting relationships between the two concepts. We see with Betty however that a mostly Analytical proof scheme alone is not sufficient to connect the relationships between mutual exclusivity and independence. This is because her conceptions of the two ideas were so powerful that she was comfortable using the four cases from her exploration to show that no relationship existed. Caroline's general proof schemes were so diverse that it is difficult to distinguish which proof scheme is be influencing her use of the definitions at any given time. From the first two cases, however, we see that little inference about how a student uses definition can be made from Harel and Sowder's (1998) proof schemes.

But we can also consider these cases with respect to Weber and Alcock's (2004) semantic and syntactic proof productions. Since he produced it immediately after changing his concept definition of mutual exclusivity to more closely resemble his concept definition of independence, we see that Alex's lemma (and therefore responses in the third interview) was a direct result of his comparing the two concept definitions. A syntactic approach to the relationship was not fruitful, however, until he changed his definition. Conversely, Betty's use of previous instantiations (a semantic approach) prevented any definition relationship between the concepts from forming. It is unclear, though, whether Betty even thought her concept definitions might need to be changed. Consider now that Caroline proved the first relationship using a examples that applied a definition of independence at the "same time," a semantic approach to the relationship. Similarly, she failed to prove the second relationship because "you would have to be putting them back if they were independent," we shows that she was relying on a specific case to try to make sense of the question. From this, we see a weak indication that a syntactic approach may play some role in aiding the adaptability of definition and that a semantic approach could be more restrictive.

From this research, we have seen how the adaptability of a student's concept image allows him or her to compare seemingly disparate concepts in new contexts. Here, the phrase “seemingly disparate” reflects the understanding of the concepts from the students’ initial points of view. This action reflects Vinner’s “interplay between definition and image,” but is different in that the participants were not comparing a definition and image of a single mathematical concept, but rather two different but related images (1991, p. 70). This interplay is not addressed in his work, but yields a result similar to that of Vinner’s interplay where an adaptation of image allows one to make sense of a perceived relationship. In this case, the adaptation of two images allowed a relationship to be perceived. Conversely, in Betty’s case, rigidity restricted her perception of a relationship between independence and mutual exclusivity.

Additionally, we have seen that merely adapting one’s conception of definitions is not necessarily sufficient to produce mathematically correct relationships. In fact, adapting one’s concept image could possibly result in an understanding less coincidental with the mathematical definition of that idea, as was witnessed with Alex’s adaptation of mutual exclusivity. Although this helped Alex to describe a relationship between the terms that was similar to the actual relationship, this is likely not the typical case. It would seem that, ideally, the student should adapt his or her concept definition toward the mathematical definition as well as adapt his or her concept image to support this definition.

**The Temporal Conception of Independence**

One interesting result of this study is the participants’ conception of independence. All three participants thought about various aspects of the two sample spaces in similar ways. When considering independence in the first sample space, each participant made two statements, synthesized here as “the dice don’t really matter” and “it does matter whether you put the card back.” Furthermore, in the second sample space, each participant made an assertion equivalent to, “all events are independent.” These statements reflect an aspect of the participants’ understanding of independence as defined on one action occurring chronologically before the other. Because of this, independence was not an effect of the events in consideration, but rather an effect of the sample space and how the actions of successfully completing the events were conducted.

It is important to note that each participant was able to correctly identify and operate with independent events in certain circumstances. For instance, when calculating the probability of an event in the first sample space, each participant could correctly identify events in the two "subsample" spaces as independent from each other as well as correctly calculate the probability of these two "subevents" happening simultaneously, and thus correctly calculate the probability of the event. From their explanations, the physical distinction and separation of these sample spaces is responsible for their independence. This was evident as the participants would hold both hands in front of them and wave the left hand when talking about one subsample space and the right hand when discussing the other. This behavior shows us that the participants understood that two different sample spaces are independent of each other, regardless of whether the participants identified them as independent.

The participants' behavior within each subsample space and, in turn, when considering entire events in the sample spaces reflected an emphasis on one event successfully occurring, time passing, and a second event successfully occurring. We call this aspect of a person's understanding of independence the "temporal conception." With the temporal conception, we see two types of sample spaces: with-memory and without-memory. With-memory sample spaces require an understanding of what happens "between the first and second event." This was reflected by the students' emphasis on whether the "first card gets replaced." Without memory sample spaces contain only independent events (i.e. the spinner, die, or coin). Since no object is removed from the whole set, probabilities never change and therefore no event can change the probability of another.

In comparing the temporal conception to Manage and Scariano's (2010) work, we consider what types of responses would result from a temporal conception of independence and/or mutual exclusivity. As we saw in this study, a temporal conception of both concepts like in Alex's case can result in a correct, albeit misleading, response. Betty's responses that the relationships are inconclusive, however, reflect a temporal conception of independence, but not of mutual exclusivity. Caroline was able to think about independence both with and without a temporal conception. We saw that with the temporal conception, Caroline determined that the two cases of with- and without-memory sample spaces would yield different responses to the first question; this was similar to Betty's reasoning using the temporal conception. But without the temporal conception, Caroline identified the correct relationship in the first direction. In the other direction (when the two events were assumed independent), Caroline was less ready to think of independence as simultaneous events. Her response that a card would have to be replaced reflected that she was again thinking of independence temporally.

From these responses, we can see that a temporal conception of both concepts (as with Alex) could lead to a "false positive" and that a temporal conception of independence but not of mutual exclusivity could lead to the third and/or fourth responses. In the latter case, instantiations of the four possible combinations of mutual exclusivity and independence led to these conclusions. This shows that a temporal conception, combined with semantic proof production, can yield a logical (from the student's perspective) response that there is not enough information to assert a relationship between independence and mutual exclusivity. From this research, though, the temporal conception does not help explain why such a large percentage of students would incorrectly respond that mutual exclusivity implies independence. The reasoning provided by Manage and Scariano that, "the issue here is the misunderstanding that 'independence' means 'separation'" provides a reasonable explanation for the high percentage of this perceived relationship (2010, p. 18). This explanation refers to the seemingly synonymous meanings of independence and mutual exclusivity in day-to-day usage.

**Contribution**

While current literature provides extensive background for proof and definition, this discussion intersects in few places (Edwards & Ward, 2004). This research has helped further research into the connections between definition and proof from three perspectives: 1) it provides a brief outline of the current literature in both areas, pointing out connections between them; 2) it uses specific cases as a context to relate definition and proof; and 3) it informs methods and approach for future research. Also, this research provides examples of students' temporal conception of independence, which was not discussed in the solicited literature. This suggests the need for possible future research specifically designed at understanding a temporal conception of independence, how to address it pedagogically, and how temporal reasoning is manifested in other mathematical areas (e.g. thinking of plotting an irrational number as a temporal process wherein one constantly approaches the value). Additionally, since this research focused on Harel and Sowder's (1998) proof schemes and definition, future research could be conducted investigating relationships between semantic and syntactic proof production and how this affects their use of definitions.

**Limitations**

In considering the limitations of this study, we must note the small number of participants. The results of this study provide what is most likely a mere glance into the nuances and intricacies of the relationships between proof and definition. This limitation was manifested in the limited exposure to the participants' proof schemes as well as the sole context of relating mutual exclusivity and independence. Although this context provided interesting and diverse results given the sample size, other mathematical relationships could have provided a broader understanding of how each participants' proof and definition related (i.e. had the participants started the discussion with concept definitions more closely related to the mathematical definitions).

From previous research we saw different methods for assessing students' proof schemes. Perhaps conducting interviews wherein the participants were asked to prove mathematical relationships on their own could have reconciled perceived inconsistencies in the participants' reasoning (i.e. Caroline's refutation of two Inductive proofs, but acceptance of a third). Participants' proof production in the first interview might also have given the researcher a better context for the participants' approaches to proof in the third interview. We must also consider the inherent differences between accepting, verifying, and producing proofs since they demand three different levels of thinking.

Another aspect of the research that could be improved is the selection, presentation, and discussion of the sample spaces in the second interview. Most of the sample spaces explored in the interviews consisted of sets not typically found in everyday events. These types of sample spaces might seem a bit contrived. There is also a matter of the way in which sample spaces were discussed in the interviews. Since a sample space is a collection of all possible outcomes on a set, it is important to clearly define what "an outcome" means. For instance, in the interviews little or no emphasis was placed on the italicized part of the phrase "*one draw of a card and one roll of a die* such that [x] card is drawn and the [y] face turns up on the die." Rather the emphasis was place on clearly stating the elements of each set that satisfied a certain quality (the second half of the previous quote). This could have caused confusion of how the events could be considered. It should be noted, however, that Alex and Betty each gave an example of independence (before the discussion of the two sample spaces) that conveyed a temporal conception. This at least shows that the problem statement did not *cause* the temporal conception.

## References


Alcock, L., & Simpson, A. (2002). Definitions: Dealing with categories mathematically. *For the Learning of Mathematics*, 22(2), 28-34.

Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *Journal of Mathematical Behavior, 24,* 125–134.

Alibert, D. & Thomas, M. (2002). Research on mathematical proof. In D. Tall (ed.), *Advanced Mathematical Thinking* (pp. 215-230). Netherlands: Scribner.

Ball, D., Hoyles, C., Jahnke, H., & Movshovitz-Hadar, N. (August, 2002). *The teaching of proof.* Paper presented at the International Congress of Mathematicians, Beijing, China.

D'Amelio, A. (2009). Undergraduate student difficulties with independent and mutually excluive events concepts. *The Montana Mathematics Enthusiast*, *6*(1&2), 47-56.

Edwards, B. S., & Ward, M. B. (2004). Surprises from mathematics education research: Student (mis)use of mathematical definitions. *The American Mathematical Monthly*, *111*(5), 411–424.

Harel, G. & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In E. Dubinsky, A. Schoenfeld and J. Kaput (eds.), *Research in Collegiate Mathematics Education*, III, 234-283. Providence, RI: American Mathematical Society.

Harel, G., & Sowder, L (2007). Toward a comprehensive perspective on teaching and learning of proof, In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning, National Council of Teachers of Mathematics*. Greenwich, CT: Information Age Publishers, Inc.

Housman, D., & Porter, M. (2003). Proof schemes and learning strategies of above-average mathematics students. *Educational Studies in Mathematics*, *53*(2), 139–158.

Jaffe, A., & Quinn, F. (1993). “Theoretical mathematics”: Toward a cultural synthesis of mathematics and theoretical physics. *Bulletin of the American Mathematical Society*, *29*(1), 1-13.

Keeler, C., & Steinhorst, K. (2001). A new approach to learning probability in the first statistics course. *Journal of Statistics Education*, *9*(3).

Kelly, I. W., & Zwiers, F. W. (1986). "Mutually exclusive and independence: unraveling basic misconceptions in probability theory." *Proceedings of the second international conference on teaching statistics* R. Davidson and J. Swift, (eds.). Victoria, BC: University of Victoria, 1-26.

Knuth, E. J. (2002). Teacher’s conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education*, *5*, 61-88.

Ko, Y. Y. (2010). Mathematics teachers’ conceptions of proof: Implications for educational research. *International Journal of Science and Mathematics Education*, *8*(6), 1109-1129.

Landau, S. I. (2001). Dictionaries: The art and craft of lexography. New York: Scribner.

Manage, A. & Scariano, S. (2010). A classroom note on: Student misconceptions regarding probabilistic independence vs. mutually exclusivity”, *Journal of Mathematics and Computer Education, 44*(1)*,* 14-20.

Martin, G. & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for research in mathematics education*, 66-80. Washington DC: NCTM.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Nickerson &Rasmussen (2009). Enculturation to proof: A pragmatic and theoretical investigation. In Lin, F. L., Hsieh, F. J., Hanna, G., & de Villiers, M. (eds.) *Conference proceedings of the international commission on mathematical instruction study 19*, Vol. 2. Taipei, Author.

Piaget, J. (1970). *Structuralism*. New York: Harper & Row.

Schoenfeld, A. (1994). What do we know about mathematics curricula? *Journal of Mathematical Behavior, 13*(1), 55–80.

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.

Shaughnessy, J. M. (1992). Research on probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematical teaching and learning* (pp.465–499). New York: Macmillan.

Tarr, J., & Lannin, J. (2005). How can teachers build notions of conditional probability and independence. In G. Jones (Ed.), *Exploring probability in school: challenges for teaching and learning* (pp. 215-238). United States: Springer.

Thurston, W. (1994). On proof and progress in mathematics. *For the Learning of Mathematics, 15*(1), 29–37.

Vinner, S. (1991). The role of definitions in the teaching and learning of mathematics. In D. Tall (Ed.), *Advanced mathematical thinking* (65-79). Dordrecht, Netherlands: Kluwer.

von Glasersfeld, E. (1995) *Radical constructivism: a way of knowing and learning.* London, UK: Falmer.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*(1), 101–119.

Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior, 23,* 115–133.

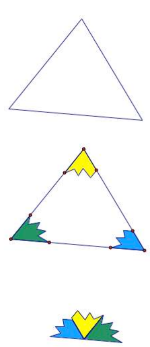
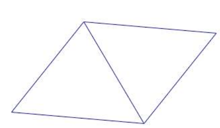
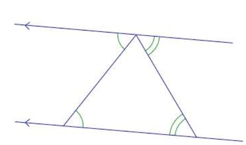
Weber, K. (2009). Mathematics majors' evaluation of mathematical arguments and their conceptions of proof. In *Proceedings for the Twelfth Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education*. Retrieved from http://mathed.asu.edu/%E2%80%8Ccrume2009/proceedings.html

Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics, 56,* 209-234.

Wu, H. (1996). The role of Euclidean geometry in high school. *Journal of Mathematical Behavior*, *15*, 221-237.

# Appendices

## Appendix A- Matrix of Proofs

| Students shown each proof and asked to evaluate/comment on mathematical soundness of each. | | |
|---|---|---|
| **Harel and Martin 1)**<br>Claim: If the sum of the digits of a whole number is divisible by three, then the number is divisible by three.<br>Proof: We randomly picked 721234182 and found that 7+2+1+2+3+4+1+8+2=30 which is divisible by three and 721234182 is divisible by three. Since this number was randomly chosen and the claim holds, then the claim is true. | **Harel and Martin 1)**<br>Claim: If the sum of the digits of a whole number is divisible by three, then the number is divisible by three.<br>Proof: Let d be an *n*-digit integer such that d = $\sum_{i=1}^{n} 10^{i-1} \cdot x_i$, where $x_i$ is the $i^{\text{th}}$ digit of *d*. Consider $10^k$ for some $1 \le k \le n$.<br>$10^k = 10^k - 1 + 1 = 9 \cdot p_k + 1$<br>Here, $p_k$ is the number with *k* digits, where every digit is 1. (and $p_0 = 1$)<br>So,<br>d = $\sum_{i=1}^{n} 10^{i-1} \cdot x_i = 9 \cdot \sum_{i=1}^{n} p_{i-1} \cdot x_i + \sum_{i=1}^{n} x_i$ .<br>From this, since 3 divides 9 times the first sum on the right, if 3 divides the sum of the digits of *d*, then 3 divides the right and so 3 divides *d*. | **Harel and Martin 1)**<br>Claim: If the sum of the digits of a whole number is divisible by three, then the number is divisible by three.<br>Proof: We show this by induction.<br>*Base case*: 12 is the smallest integer with more than one digit that is divisible by 3. 1+2 = 3.<br>*Induction step*: Assume true for all multiples of 3 less than or equal to *n*, a multiple of 3. Consider *n*+3, the next multiple of 3. Since the sum of the digits of *n* is divisible by 3, this sum plus three is also divisible by three. Therefore, claim holds. |
| **Harel and Martin 2)**<br>Claim: If a divides b and b divides c, then a divides c.<br>Proof: a divides b; this means there is some number k, such that k*a=b. Also, b divides c, which means there is a number n, such that b*n=c. Now substitute for b in the last equation, and we get n*(k*a)=c. By the associative property, (n*k)*a = c. Therefore a divides c. | **Harel and Martin 2)**<br>Claim: If a divides b and b divides c, then a divides c.<br>Proof: Let's pick any three numbers, taking care that the first divides the second, and the second divides the third; 49 divides 98, and 98 divides 1176. Does 49 divide 1176? The answer is yes. Since these numbers were chosen randomly and the claim holds, then the claim is true. | **Harel and Martin 2)**<br>Claim: If a divides b and b divides c, then a divides c.<br>Proof: 3 does not divide 5, and 5 does not divide 7. We see that<br>3 does not divide 7. On the other hand, 3 divides 6, and<br>6 divides 12. In this case, 3 also divides 12. |
| **Plaxco 1)**<br>Claim: The sum of the interior angles of a triangle is 180°.<br>Proof: We see by cutting each corner of the triangle and rearranging the angles that the sum of the interior angles is a straight line, which equals 180°.<br> | **Plaxco 1)**<br>Claim: The sum of the interior angles of a triangle is 180°.<br>Proof: A triangle is one half of a quadrilateral, which has a sum of interior angles equal to 360°.<br> | **Plaxco 1)**<br>Claim: The sum of the interior angles of a triangle is 180°.<br>Proof: Drawing a line through a vertex and parallel to that vertex's opposite side, we see the triangle forms two transversals of parallel lines. Since alternate interior angles are congruent, the constructed line (180°) is the sum of the triangle's interior angles.<br> |

**Appendix B- Interview 2 Sample Spaces**

| | |
|---|---|
| Sample Space 1 | Consists of:<br>1 deck of cards and 1 six-sided die.<br>Tasks:<br>Identify pairs of events as ME, IND, both, or neither. Discuss methods for each answer.<br>Events:<br>A= "Drawing a spade and rolling any number."<br>B= "Drawing any card and rolling a 3."<br>C= "Drawing a spade and rolling a 3."<br>D= "Drawing a heart and rolling a 4." |
| Sample Space 2 | Consists of:<br>A spinner in five sections<br>(Blue: 30%, Red: 20%, Green: 20%, Orange: 20%, Yellow: 10%) and a dime.<br>Tasks: Identify pairs of events as ME, IND, both, or neither. Discuss methods for each answer.<br>Events:<br>A= "Spinning a blue and landing heads"<br>B= "Spinning a red and landing tails"<br>C= "Spinning a green and landing heads"<br>D= "Spinning a blue and landing tails"<br>E= "Spinning a yellow or blue and landing heads"<br>F= "Spinning a purple and landing on heads" |

**Appendix C- IRB Permission Form**

# VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

## Informed Consent for Participants of a Research Project Involving Human Subjects

*Title of Project*: Relationship Between Students' Proof Schemes and Definitions

*Investigators*: David Plaxco; Anderson Norton

I. **Purpose of this Research/Project:** To improve knowledge about students' understanding and misconceptions about mathematical definitions and how this relates to their general proof scheme.

II. **Procedures:** Study participants will be interviewed three times with each interview lasting approximately 40-60 minutes.

The interview will be recorded on a digital audio recorder and video recorder and notes will be taken during the interview. I will collect your work and use it when reflecting on the interview.

Please feel free to respond to the questions and/or tasks in these interviews to a degree with which you are comfortable. These questions are meant to give me a better understanding of how people think about mathematics. You will not be graded for or against credit in any class. Your responses will not be associated with your name except to Dr. Anderson Norton and David Plaxco. This includes your professors or any person who might read the findings of this study.

After each interview, I will type a transcript of the interview. All transcripts of the interviews will be stored on my computer in a password-protected file. After each transcript is typed, I will contact you in order to have you check the transcript. You can ask me to remove any part of the transcript that you would like removed.

III. **Risks to Confidentiality:** This paragraph is to inform you that the access to transcripts of the interviews will only be allowed to primary investigator and co-investigator. The faculty/staff and students of Virginia Tech's Mathematics Department will not see your individual responses or know your identity. Pseudonyms will be used to identify you in all interviews and written materials. You may decline to answer any questions that you don't feel comfortable with. All consent forms and data resulting from this study will be kept in a locked file cabinet on campus and all the data will be digitized and stored on a secure computer network. All data will be coded with labels and numbers. No personal identifying marks will be present on any data forms. Data will be analyzed without personal identification.

**III.** **Benefits and Compensation:** You will receive no compensation for you participation in this study. However, identifying themes and patterns can inform the mathematical community on understanding of undergraduates' concepts of proof.

**IV.** **Freedom to Withdraw:** Participants are free to cease the participation at any time without prejudice, penalty, or any other negative consequence.

**V.** **Subject's Responsibilities:** I voluntarily agree to participate in this study. I have the following responsibilities:

- I agree to answer questions honestly.
  Initial ________
- I agree to allow the researcher to record the interview on digital recording device. Initial ________
- I agree to allow the researcher to use a non-identifying direct quote.
  Initial ________

**VI.** **Participant's Permission:** I have read the Consent Form and conditions of this research project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent:

_______________________________________________ Date__________

Participant signature

Thank you for your participation in this research project. It is my hope that your cooperation will contribute to the mathematics community in a positive way.

David Plaxco - Investigator
dplaxco@vt.edu

**[NOTE: Participants must be given a complete copy (or duplicate original) of the signed Informed Consent.]**