Two Aspects of Proof: Examining the Amount of Logic in Student-Constructed Proofs and Mathematicians' Actions in Recovering From Proving Impasses

Milos Savic

Dr. Milos Savic is an assistant professor at the University of Oklahoma. He researches in aspects of proof and proving in upper-level undergraduate courses. A Project NExT member of the MAA, he also is a co-founder of a proof research group and the Oklahoma Research in Undergraduate Mathematics Education group. He finds time at home to play with his daughter and wife. Finally, he is a proud BSU Mathematics graduate (class of 2004).

To obtain a Master's or PhD in mathematics, or even to succeed in proof-based courses in an undergraduate mathematics major, one must often be able to construct original proofs, a common difficulty for students [18, 30]. This process of proof construction is usually explicitly taught, if at all, to U.S. undergraduates as a small part of a course, such as linear algebra, whose stated goal is something else, or in a transition-to-proof or "bridge" course. Students might also get discouraged when attempting a proof, perhaps due to the differences between proving and prior exercises [15] asked of them. Students may often complain about "getting stuck." In this article, I attempt to address two questions in proving: what extent does logic appear on the surface of student-constructed proofs, and what do mathematicians do when they "get stuck."

Part I - An Examination of Logic in Student-Constructed Proofs

When universities do offer a transition-to-proof course, professors often teach some formal logic (predicate and propositional calculus) as a background for proving. But how much logic actually occurs in student-constructed proofs? I begin to answer this question by first searching for uses of logic in a "chunk-by-chunk" analysis of student-constructed proofs from a graduate "proofs course." If formal logic occurs a substantial amount, then teaching a unit on predicate and propositional calculus might be a good idea; however, if formal logic occurs infrequently, then teaching it in context, while teaching proving, may be more effective.

Background Literature

Currently, at the beginning of transition-to-proof courses, professors often include some formal logic, but how it should be taught is not so clear. In [9, p.895] Epp stated that, "I believe in presenting logic in a manner that continually links it to language and to both real world and mathematical subject matter". However, some mathematics education researchers maintain that there is a danger in relating logic too closely to the real world: "The example of 'mother and sweets' episode, for instance, which is 'logically wrong' but, on the other hand, compatible with norms of argumentation in everyday discourse, expresses the sizeable discrepancy between formal thinking and natural thinking", [4]. In the mother and sweets scenario, the mother says to the child, "If you don't eat, you won't get any sweets" and the child responds by saying, "I ate, so I deserve some sweets." Other authors have noticed that the way logic is taught in transition-to-proof courses is at variance with how it is actually used in proving: "Beginning logic courses often seem to present logic very abstractly, in essence as a form of algebra, with examples becoming a kind of applied mathematics", [22, p.8].

There are also those who think that logic does not need to be explicitly introduced at all. For example, Hanna and de Villiers [13, p.311] stated, "It remains unclear what benefit comes from teaching formal logic to students or to prospective teachers, particularly because mathematicians have readily admitted that they seldom use formal logic in their research." Selden and Selden, in [24, p. 347] claimed that "logic does not occur within proofs as often as one might expect ... [but] [w]here logic does occur within proofs, it plays an important role." Taken together, these differing views suggest that it would be useful for mathematics education researchers to further examine the role of logic and logic-like reasoning within proofs in order to inform mathematics lecturers on the ways they might best include logic in transition-to-proof courses. However, to date, only a little such research has been conducted [5].

Another interesting idea that has been expressed about proofs in general is that deduction occurs in proofs in a "systematic, step-by-step manner" [3]. In fact, one professor quoted by Ayalon and Even [3, p.34] expressed the view that a student "thinks about something, he draws a conclusion, which brings him to the next thing... Logic is the procedural, algorithmic structure of things." Rips,[19] looked at proof in a slightly more sophisticated way: "At the most general level, a formal proof is a finite sequence of sentences $(s_1, \ldots, s_k)$ in which each sentence is either a premise, an axiom of the logical system, or a sentence that follows from the preceding sentences by one of the system's rules." Instead of sentences, I partition student-constructed proofs into usually smaller "chunks" to begin to answer the question I posed earlier on logic.

Research Settings

The "proofs course," *Understanding and Constructing Proofs*, was offered at a large U.S. Southwestern state university, giving master's and PhD's in mathematics. Students in the course were first-year mathematics graduate students along with a few advanced undergraduate mathematics majors. For this course, the students were given professor-created notes with a sequence of definitions, questions, and statements of 42 theorems dealing with topics such as sets, functions, real analysis, algebra, and topology. For example, three theorems that were proved by the students were: "The product of two continuous [real] functions is continuous"; "Every semigroup has at most one minimal ideal"; and "Every compact, Hausdorff topological space is regular." The topics in the course were of less importance than its focus on the construction of differing kinds of proofs. All proofs analyzed in this study were student-constructed and verified as being correct by the professors.

Research Methodology

The 42 proofs from the "proofs course" were first subdivided into "chunks" for coding. The "chunks" are similar to those in Miller's article, [17] in which he stated that chunks are a "meaningful unit" in thinking. In the analysis described here, a chunk can refer to a sentence, a group of words, or even a single word, but always refers to a unit in a proof. During several iterations of the coding process, 13 categories, such as "Informal inference" and "Assumption," emerged.

The categories and the chunks sometimes co-emerged, that is, the categories sometimes influenced the chunking. For example, "Then $x \in A$ and $x \in B$" might have been treated as a single chunk because it arose from $x \in A \cap B$ and the definition of intersection. However, it could have been split into "Then $x \in A$" and "and $x \in B$" because the two chunks seemed to follow from separate warrants.

In this proposal, I discuss in detail just 5 of the 13 categories. The first two of these deal with the question posed at the beginning of this paper, "Where is the logic in student-constructed proofs?" The remaining three categories are those that occurred most often.

The Categories

Informal Inference (**II**) is the category that refers to a chunk of a proof that depends on common sense reasoning. While I view informal inference as being logic-like, it seems that when one uses common sense, one does so automatically and does not consciously bring to mind any formal logic. For example, given $a \in A$, one can conclude $a \in A \cup B$ by common sense reasoning, without needing to call on formal logic.

By *Formal Logic* (**FL**) in this paper I mean the conscious use of predicate or propositional calculus going beyond common sense. The distinction is that formal logic is the logic a university student does not normally possess before entering a transition-to-proof course. Modus Tollens and DeMorgan's Laws are two examples of formal logic that are usually not common sense for such students [1, 2]. For example, given $x \notin B \cup C$, one can conclude $x \notin B$ and $x \notin C$, a typical use of DeMorgan's Laws that students often do not perform automatically, or do perform automatically, but incorrectly.

Definition Of (**DEF**) refers to a chunk in a proof that calls on the definition of a mathematical term. For example, consider the line "Since $x \in A$ or $x \in B$, then $x \in A \cup B$." The conclusion "then $x \in A \cup B$" implicitly calls on the definition of union.

Assumption (**A**) is the code for a chunk that creates a mathematical object or asserts a property of an object in the proof. The category is further divided into two sub-categories: "Choice" and "Hypothesis." *Assumption (Choice)* refers to the introduction of a symbol to represent an object (often fixed, but arbitrary) about which something will be proved – but not the assumption of additional properties given in a hypothesis. In contrast, *Assumption (Hypothesis)* refers to the assumption of the hypothesis of a theorem or argument (often asserting properties of an object in the proof). An example to demonstrate the difference between the two is provided by the theorem "For all $n \in \mathbb{N}$ if $n > 5$ then $n^2 > 25$." The chunk "Let $n \in \mathbb{N}$" would be coded Assumption (Choice), and the chunk "Suppose $n > 5$" would be coded Assumption (Hypothesis).

Interior Reference (**IR**) is the category for a chunk in a proof that uses a previous chunk as a warrant for a conclusion. For example, if there were a line indicating $x \in A$ earlier in the proof, then a subsequent line stating "Since $x \in A \ldots$" later in the proof would be an interior reference.

The Results

In the chunk-by-chunk analysis of the proofs in the "proofs course," just 6.5% (44 chunks) of the 673 chunks were Informal Inference, and just 1.9% (13 chunks) were Formal Logic. However, I found that 30% (203 chunks) were Definition Of, 25% (166 chunks) were Assumption, and 16% (108 chunks) were Interior Reference.

Table 1 below shows the chunk categories, complete with the rounded percentages:

Discussion

At first glance, these results may seem surprising. While the chunk-by-chunk coding is a convenient tool for a surface analysis of a finished written proof, there are underlying structures to, and within, proofs, such as proof by contradiction. I see these as "logic-like structures" that are not often explained in the predicate and propositional calculus discussed in most transition-to-proof courses. For example, if one wishes to prove "For all $x \in A$, $P(x)$", one starts with "Let $x \in A$" and reasons towards "$P(x)$". Structuring a proof in this way has the effect of using logic.

The fact that from the "proofs course," 30% of the chunks were definitions and 25% were assumptions suggests that there is a need to teach undergraduates how to introduce mathematical objects into proofs and how to read and use definitions. Indeed, there have been documented instances of students' struggle with definitions [8]. Another implication for teaching that stems from this research is that because formal logic occurs fairly rarely, one might be able to teach it in context as the need arises.

	"Proofs class"	% of chunks
# of chunks	673	
A	166	24.7
ALG	23	3.3
C	56	8.3
CONT	4	0.6
D	32	4.7
DEF	203	30.2
ER	17	2.5
FL	13	1.9
II	44	6.5
IR	108	16
REL	3	0.4
SIM	3	0.4
SI	4	0.6

Table 1: Distribution of the "proof chunks" by category.

Future Research

It would be interesting to examine whether the kinds of chunks used in proofs varies by mathematical subject area. For example, would topology have a different distribution of categories of chunks than abstract algebra? Indeed, several mathematics professors have suggested that I code chapters of various textbooks to see how much formal logic occurs in them. Also, it may be that the kind of formal logic taught explicitly at the beginning of many transition-to-proof courses is actually psychologically, and practically, disconnected from the process of proving for students. This disconnect might lead to future difficulties in many of the proof-based courses in students' subsequent undergraduate and graduate programs. An additional interesting question that arises from the "proofs course" itself is: How many beginning graduate students need a course specifically devoted to improving their proving skills?

In future research, one might also look for instances of logic-like structures and techniques in student-constructed proofs. Solow [26] and Velleman [28] both discuss logic-like structures and techniques for proving, but many other transition-to-proof books touch on this only very briefly, if at all. Can one identify a range of logic-like structures that students most often need in constructing proofs? Further, one might investigate the degree of a prover's automated behavioral knowledge of logic-like structures that could help reduce the burden on his or her working memory. This might free resources to devote to the problem-solving aspects of proofs. That this might be the case was suggested by Selden, McKee, and Selden [23].

Finally, there may be additional logic that does not appear in a final written proof, but that might occur in the actions of the proving process. This would be interesting to investigate.

Part II - What Do Mathematicians Do When They Encounter an Impasse?

In examining mathematicians' proof construction practices, the study reported here focused on impasses and how the mathematicians overcame those impasses, including incubation and the resulting insight. An impasse is colloquially referred as "getting stuck" or "spinning one's wheels." These ideas have been examined in the psychology and mathematics education literatures, mainly in analyzing problem solving, but there has been little research on them during proving. A brief discussion of this literature provides background for my somewhat different use of these terms in examining and analyzing proof construction.

Background Literature

In the psychology literature, Duncker [7] defined an impasse as a mental block against using an object in a new way that is required to solve a problem. One way problem solvers sometimes recover from impasses is through incubation. Incubation, according to Wallas [29], is the process by which the mind goes about solving a problem, subconsciously and automatically. It is the second of Wallas' four stages of creativity, which are:

- preparation (thoroughly understanding the problem),
- incubation (when the mind goes about solving a problem subconsciously and automatically),
- illumination (internally generating an idea after the incubation process), and
- verification (determining whether that idea is correct).

While psychologists' treatments of impasses, incubation, and insight may be useful in investigating a number of instances of creativity and problem solving, especially simpler instances, analyzing the construction of original proofs in mathematics seems to call for some modification of them. For example, all of the 117 experiments considered by Sio and Ormerod [24] in their meta-analysis of incubation studies used incubation periods of just 1-60 minutes. Mathematicians routinely take much more time to overcome impasses in their research, and their proofs tend to be considerably longer and more complex.

Problem solving and incubation in mathematics and mathematics education

To date, research on problem solving and incubation in the mathematics education literature has been sparse and primarily anecdotal. Creativity and incubation are rarely captured in research: "[S]tudying a mathematician's or student's creativity is a very difficult enterprise because most traditional operationalized instruments fail to capture extra cognitive traits, such as beliefs, aesthetics, intuitions, intellectual values, self-imposed subjective norms, spontaneity, perseverance standards, and chance" (Freiman & Sriraman, [10, p. 23]). Some instruments that have been used to capture creativity or incubation in mathematics education include video interviews, written work, or problem/proving sessions in front of a camera, [11].

In his investigation of mathematicians' practices, Hadamard [12] mailed surveys to mathematicians around the world to collect information on what mathematicians do. Nicolle, one of the mathematicians that responded to his survey, concluded that "contrary to progressive acquirements, such an act [discovery of a solution after an impasse] owes nothing to logic or to reason. The act of discovery is an accident," [12, p.19]. While discovery may be an "accident," the actions taken to allow the mind to have such an accident might be quite deliberate.

How incubation can help mathematicians is still somewhat of a mystery. Sriraman [27, p.30] conjectured that "the mind throws out fragments (ideas) which are products of past experience." Those fragments are brought to light as insights. Hadamard, Liljedahl, and Sriraman uncovered, through interviews with mathematicians, some evidence that mathematicians use incubation and then experience insights when solving problems. I hope to add to this literature, partly by narrowing the focus to theorem proving, making observations in a realistic setting, and supplying notes on an unfamiliar, but accessible, algebraic topic for mathematicians to work on. I will also focus on describing reportable or observable events, as opposed to speculating on complex mental activity outside of consciousness.

Research Questions

According to the research described above, the creative process often includes impasses and incubation, but previous researchers have had trouble capturing the proving process in a realistic setting in real time (Liljedahl, [15]; Freiman, and Sriraman, [10]). Also, studying mathematicians could help with teaching students (Weber, [31]). In this vein, I investigated the following two questions:

1. Can researchers capture much of the proving process without time and place constraints?

2. What do mathematicians do when they reach a proving impasse and is what they do helpful?

Impasses and Incubation in Proving

The above meanings for impasse and incubation used in describing results in psychology seem not to "fit" the proving process well. Thus I introduce somewhat altered meanings.

By a (proving) *impasse*, I mean a period of time during the proving process when a prover feels or recognizes that his or her argument has not been progressing fruitfully and that he or she has no new ideas. What matters is not the exact length of time, or the discovery of an error, but the prover's awareness that the argument has not been progressing and requires a new direction or new ideas. Mathematicians themselves often colloquially refer to impasses as "being stuck" or "spinning one's wheels." This is different from simply "changing directions," when a prover decides, without much hesitation, to use a different method, strategy, or key idea, and the argument continues.

I mean by *incubation* a period of time, following an attempt to construct at least part of a proof, during which similar activity does not occur. After incubation, a prover might have an insight, that is, the generation of a new idea. That insight might be helpful, and it might move the argument forward. For some of the major incubations described here, resulting insights occurred, they were helpful, and they moved the argument forward. However, in future studies, in which the participants might be less skilled or the proofs might be more difficult, incubation might be less fruitful. Also, all but one of the major incubations described here were purposeful, but with future studies in mind, I do not include purposefulness in the meaning of incubation. A long proving process might entail several impasses and a number of incubation periods (and subsequent insights), only some of which ultimately contribute to the final proof.

Research Methodology

Nine PhD mathematicians (three algebraists, three topologists, two analysts, and one logician) agreed to participate in this study on proving. All mathematicians were given pseudonyms, Drs. A-I, in order of participation. The mathematicians were selected by the author to participate based on convenience and rapport (seven of the nine were the author's past professors). All were from one southwestern PhD-granting university and eight of the nine mathematicians were currently active in their research. They were provided with a set of slightly modified notes (Savic, [20]; Selden, McKee, and Selden, [23]) on semigroups, containing ten definitions, seven requests for examples, four questions to answer, and 13 theorems to prove. The topic, semigroups, was selected because the mathematicians would hopefully find the material easily accessible, and because there are two theorems towards the end of the notes (Theorems 1 and 2 of the Appendix) that have caused substantial difficulties for beginning graduate and upper-level undergraduate students. During their exit interviews, two mathematicians offered that the choice of semigroups had been judicious, because they had been able to grasp the definitions and concepts quickly, and because at least one of the theorems had been somewhat challenging to prove.

The data collection began with four mathematicians writing proofs on tablet PCs, and later five mathematicians wrote their proofs with a LiveScribe pen and special paper. The switch from tablet PC to LiveScribe pen was done for several reasons. First, tablet PCs are relatively more expensive than LiveScribe pens and the corresponding paper. Second, the size of a "movie file" for a tablet PC screen capture of 16 minutes is one gigabyte, whereas an almost five hour proving session on a LiveScribe pen is just 60 megabytes. Third, the mathematicians were much more comfortable with pen and paper than with the tablet PC and stylus. Fourth, there were no visual or auditory quality differences between the data collected using the two techniques. This allowed for a smooth transition of data collection techniques to one that I felt was the most comfortable for the participants, and provided all the real-time data collection that I needed. One or two days after this initial analysis of each mathematician's work, I conducted an exit interview, during which I asked about their proofs and proof-writing. I also conducted two separate focus groups with both the Tablet PC group and the LiveScribe pen group of mathematicians where they could discuss and reminisce about the notes and their work.

Results

Summary data

Four of the nine mathematicians that participated in the study had difficulty with the technology and thus did not produce "live" data. Difficulties included not loading certain programs correctly, not remembering to press the record button, and computer error with installing CamStudio software. However, all four provided good written data, whether it was with the tablet PC on OneNote or with writing on the LiveScribe paper without audio/video recording. From this data I could still conclude that some mathematicians had impasses because they were candid in writing all their work, including crossing out failed attempts, while also providing descriptions of their thought processes (when prompted) during the follow-up interview.

The average total work time on the technology per mathematician was two hours and five minutes. This time was calculated by adding the durations of their actual work, obtained from the date and time stamps. The average time from the first "clocked in" time-and-date stamp until the last "clocked out" time-and-date stamp was 19 hours, 56 minutes. The average number of pages written was slightly under 13. These three statistics allow one to conclude that the mathematicians expended considerable effort on proving the theorems and producing examples. Six of the nine mathematicians had impasses when proving one of the final two theorems (Theorems 1 and 2 of the Appendix). Most mathematicians correctly proved most of the theorems very quickly until they got to those final two theorems. Some actions to overcome impasses were common with most mathematicians in the study, other actions were unique to those individuals.

Case study of a mathematician's proving process

Below are descriptions of an impasse, an incubation period, and an insight leading to a proof for Dr. A, an applied analyst. Dr. A used a tablet PC and part of his work is described below using the time-and-date stamps.

Dr. A

In proving Theorem 2 of the Appendix, namely "If S is a commutative semigroup with minimal ideal K, then K is a group," Dr. A experienced an impasse, an incubation period, and a resulting insight – these are indicated in bold in Table 2 of the Appendix. The abbreviated, interpreted timeline in Table 2 of the Appendix illustrates this.

Dr. A indicated in his exit interview where he had had an impasse, noting "One has to show there aren't any sub-ideals of the minimal ideal itself, considered as a semigroup, and that's where I got a little bit stuck." This is because the concept of ideal really depends on the containing semigroup, here S or K. Dr. A also indicated how he deliberately generally recovers from impasses: he prefers to get "un-stuck" by walking around, but distractions caused by his departmental duties also help. That is, he often takes a break from his creative work by purposely doing something unrelated. In this case, Dr. A took several such breaks (e.g., 11:18 AM – 11:32 AM, 12:01 PM – 12:22 PM, 12:28 PM – 12:55 PM) to prove one proof, but only the last one yielded a useful new idea.

Mathematicians' Actions to Recover from Impasses

From an analysis of data from Tablet PC or LiveScribe pens, exit interviews, and the focus group sessions, I was able to gather actions that the participating mathematicians used to recover from impasses. Some actions to recover from an impasse were observed in the proving processes of the mathematicians in the data collected while they worked alone, whereas other actions were first mentioned during the exit interviews or focus group discussions. Most of the actions that the mathematicians took to overcome their proving impasses were enacted more or less automatically and were not mentioned during their proving sessions. However, the mathematicians did acknowledge these actions either during their exit interviews or during the focus group discussions.

Impasse recovery actions that use mathematics

(a) *Using methods that occurred earlier in the proving session*: Some of the mathematicians in this study tried to use a proving technique that they had used earlier in the proving session to overcome an impasse.

"It would be fairly easy to prove ⋯ it's likely an argument, kind of like the one I already used ⋯." (Dr. H)

(b) *Using prior knowledge from their own research*: There were mathematicians in this study who tried to use ideas from their own research to overcome an impasse.

"I'm trying to think if there's anything in the work that I do that ⋯ I mean some of the stuff I've done about subspaces of, umm ⋯ there are things called principal shift invariance spaces that the word principal comes into play." (Dr. A)

(c) *Using a (mental) database of proving techniques*: One of the mathematicians, Dr. F, stated that she had a (mental) database of proving techniques in her head.

"Your brain is randomly running through arguments you've seen in the past ⋯ standard techniques that keep running through my head, sort of like downloading a whole bunch at the same time and figuring out which way to go." (Dr. F)

(d) *Doing other problems in the problem set and coming back to the impasse*: Five of the nine mathematicians in the study approached their proving impasses by moving forward to consider the rest of the problems in the notes.

"I moved on because I was stuck ⋯ maybe I was going to use one of those examples [Question 22] ⋯ I might get more information by going ahead." (Dr. B)

(e) *Generating examples or counterexamples*: Three of the mathematicians in the study attempted to construct counterexamples to some of the theorems when they felt they had not been correctly stated.

"At first I thought, 'How could I prove [Theorem 20]?' And I didn't immediately think of a proof. Then I thought, 'what about a counterexample?' and pretty quickly I came up with a counterexample, of course which turns out *not* to be right." (Dr. G)

(f) *Doing other mathematics*: Some mathematicians indicated that they might go to another project to help them overcome proving impasses.

"What I try to do is to keep three projects going ⋯ I make them in different areas and different difficulty levels ⋯." (Dr. E)

Impasse recovery actions that are non-mathematical

(a) *Taking a break*: Some mathematicians indicated that sometimes they may choose to walk around to overcome a proving impasse. This action is the first listed that is non-mathematical.

"When I'm stuck, I often feel like taking a break. And indeed, you come back later and certainly for a mathematician you go off on a walk and you think about it." (Dr. G)

(b) *Doing tasks unrelated to mathematics:* This is the second non-mathematical action unrelated to an impasse. This action was also perhaps the most unusual, and Dr. E seemed slightly embarrassed when he reported the action to me.

"Yeah I'll do something else, and I'll just do it, and if there's a spot where I get stuck or something, I'll put it down and I'll watch TV, I'll watch the football game, or whatever it is, and then at the commercial I'll think about it and say, 'yeah that'll work' ⋯." (Dr. E)

(c) *Going to lunch/eating:* This action was shown to be effective above with Dr. B.

"So I had spent probably the last 30 minutes to an hour on that time period working on number [Theorem] 1 going in the wrong direction. Ok, so I went to lunch, came back, and while I was at lunch, I wasn't writing or doing things, but I was just standing in line somewhere and it [an insight] occurred to me the ⋯ (laughs) ⋯ how to solve the problem." (Dr. B)

(d) *Sleeping on it:* The final action to overcome an impasse seems to be the easiest for a mathematician. Proving can involve mental exhaustion, so resting can help one's exploration for new ideas.

"It often comes to me in the shower ⋯ you know you wake up, and your brain starts working and somehow it [an insight] just comes to me. I've definitely gotten a lot of ideas just waking up and saying 'That's how I'm going to do this problem'." (Dr. F)

Advantages of Incubation

A majority of the nine mathematicians in this study exhibited impasses and recoveries from those impasses, including some due to incubation. Furthermore, there were a number of instances in which impasses and recoveries from them, that is, incubations, might have occurred in a way that could not be easily observed. For example, all of the mathematicians reported that when they first received the notes they immediately read them to estimate how long the proofs might take, but none started proving right away. In addition, there were periods during the proving sessions when nothing was recorded, and there were also substantial gaps between the "clock in" and "clock out" times during the proving sessions. Furthermore, when the mathematicians next "clocked in" after having left a proof attempt without finishing it, they almost always had a new idea to explore.

In the focus groups, the mathematicians also discussed methods of impasse recovery and what amounts to incubation (that can occur independent of an impasse). They all did this in a relaxed, assured way, not like someone discussing something unfamiliar, but rather like someone discussing practices built up over some time. They described a remarkable number of ways of recovering from an impasse. Furthermore, they mentioned general benefits that appear to go beyond just restarting an argument, such as clearing the mind or developing more understanding of the theorem. During one focus group interview, Dr. G stated: "When we are working on something, we are usually scribbling down on paper. When you go take a break, ⋯ you are thinking about it in your head without any visual aids ⋯ [walking around] forces me to think about it from a different point of view, and try different ways of thinking about it, often global, structural points of view."

He stated that there is no "scribbling on paper." Doing this, that is, thinking more generally, he believed, might assist in understanding the structure of a problem or even of an area of mathematics. In a somewhat similar vein, Dr. F offered the following in her exit interview: "You just come back with a fresh mind. [Before that] you're zoomed in too much and you can't see anything around it anymore." Her perspective is more of "freshness" which may allow for different ideas instead of Dr. G's "lack of visual aids" approach. But both encourage going away from the proof or problem for the generation of newer ideas. Dr. A stated, "I do have a belief that if I walk away from something and come back it's more likely that I'll have an idea than if I just sit there." These remarks indicate that some mathematicians take deliberate actions to overcome impasses and also to improve the breadth or quality of their perspectives.

Deliberate incubation has been shown in the psychology literature to result in a greater incubation effect than merely being interrupted during the problem-solving process. "Individuals who took breaks at their own discretion (a) solved more problems and (b) reached fewer impasses than interrupted individuals" (Beeftink, van Eerde, and Rutte, [6]). Ironically, interruption seems to have been useful in the case of Dr. B who said that he would have worked non-stop if he had not been interrupted for lunch with his family. This also agrees with the psychology literature: "It was also found that interrupted individuals reached fewer impasses than individuals who worked continuously on problems" (Beeftink, van Eerde, and Rutte, [6]). Finally, in their meta-analysis of the incubation literature, Sio and Ormerod, [24] stated that "low-demand tasks" done in the incubation period yielded positive incubation effects. When compared to high-demand tasks, they stated: "There remains a possibility, of course, that a sufficiently light load might allow additional covert problem solving compared with a heavier task load" (p.107). One mathematician, Dr. E, stated that he had different projects in different difficulty levels and rotated among them, which corresponds well with the positive effect Sio and Ormerod found in the incubation literature.

Education Implications

Similar experiences to the mathematicians can probably be provided to undergraduate mathematics students who are not yet familiar with constructing proofs by considering the problem-solving literature. A problem that is likely to generate impasses is probably close to what Schoenfeld, [26] has described as a "rich" problem:

- The problem needs to be accessible. That is, it is easily understood, and does not require specific knowledge to get into.
- The problem can be approached from a number of different ways.
- The problem should serve as an introduction to important mathematical ideas.
- The problem should serve as a starting point for rich mathematical exploration and lead to more good problems. (as cited by Liljedahl, [15, pp.187-188])

According to Mann, [16], "if mathematical talent is to be discovered and developed, changes in classroom practices and curricular materials are necessary. These changes will only be effective if creativity in mathematics is allowed to be part of the educational experience" (p.237). Using Schoenfeld's "rich" problems, a teacher might introduce a mathematical experience quite like a mathematician, or perhaps an AHA! moment, to a student.

Furthermore, students may need to experience successes in order to acquire confidence in their proving ability, and telling them what mathematicians do when they "get stuck" might help them when they have "no idea what to do next." Moreover, there is support from the psychology literature about the positive effects of incubation in the classroom. Sio and Ormerod,[24] cited four articles where "educational researchers have tried to introduce incubation periods in classroom activity, and positive incubation effects in fostering students' creativity have been reported" (p.94).

Future Research

Using LiveScribe pens and the corresponding paper provides a naturalistic setting for provers while gathering real-time data. Since one can observe what a mathematician does during the proving process, those same techniques might be used with undergraduate mathematics students in a transition-to-proof or proof-based course. In particular, how can we use this data collection technique in the classroom? Would it benefit students to have LiveScribe pens with which to do their homework so that teachers can analyze their proving processes? This study gave the mathematicians unlimited time for proving the theorems in the notes. However, some "breaks" could have been because of other factors instead of coming to an impasse. How can we gain additional information on when and how incubation benefits mathematicians or students? Finally, mathematicians seem to know in some cases that they need to take a break for generating ideas. How can we encourage students to take some of the actions demonstrated by the mathematicians in this study to recover from their proving and problem-solving impasses?

Appendix

Theorem 1. *If S is a commutative semigroup with no proper ideals, then S is a group.*

Theorem 2. *If K is a minimal ideal of a commutative semigroup S, then K is a group.*

7/13/11 3:48PM 9 min	At this time Dr. A first attempted a proof of Theorem 21. He stopped and moved on to Question 22.
Break 3:57 PM - 4:01 PM	
7/13/11 4:01PM 16 min	Continuing later, when he had finished Question 22, Dr. A scrolled up to his first proof attempt. He looked at his answer to Question 22, and at the ten minute mark, erased his first proof attempt. He then scrolled back to his proof of Theorem 20, viewed it for one minute, and wrote "the argument above proves that K has a multiplicative identity in S." There was a brief pause, after which he scrolled up to the proof of Theorem 20 again for the final 30 seconds. Proving ended for the day at 4:17.
Break 4:17 PM, July 13th - 11:07 AM, July 14th	
7/14/11 11:07 AM 11 min	The next day Dr. A again started attempting to prove Theorem 21. But this time he used a mapping ϕ^{-1}that multiplied each element by a fixed k_0 (an idea from his own research). He struggled with some computations until the end of this "clocked in" period.
Break 11:18 AM - 11:32 AM	
7/14/11 11:32 AM 5 min	When he "clocked in" again, Dr. A again worked with the mapping idea and then wrote "I don't know how to prove that K itself is a group. For example,I don't know how to show that there is an element of K that fixes k_0," acknowledging that he was at an **impasse**.
7/14/11 11:38 AM 23 min	However, Dr. A continued trying unsuccessfully to use his mapping idea.
Break 12:01 PM - 12:22 PM	
7/14/11 12:22 PM 6 min	When Dr. A "clocked in" again, he continued trying unsuccessfully to use his mapping idea. For example, he wrote, "To prove ϕ is well-defined, let $tk_0 = x, tk_1 = k_2$. Let v be any other element of S such that $vk_0 = x$. Choose any $w \in S$ s.t. $wx = k_2$. Then $vk_1 = vwx = vwtk_0 = twx = tk_1$. So $\phi(t)$ is determined once tk_0 is determined."
Break 12:28 PM - 12:55 PM	
7/14/11 12:55 PM 5 min	Later on, when he "clocked in" again, after a 33-minute gap (which might be considered an **incubation** period), Dr. A proved Theorem 21 writing "Proof of theorem: We just need to show that K itself has no proper subideals. But K is principally generated, i.e., fix any $k_0 \in K$ and $K = \{sk_0 : s \in S\}$ since K is [a] minimal [ideal]. If L were a proper ideal of K . . ." Notice that this idea (an **insight**) for proving Theorem 21 differs from the idea he had tried 33 minutes earlier.

Table 2: Timeline for Dr. A showing an impasse and an insight

References

[1] J. Anderson, Cognitive Psychology and its Implications. San Francisco: Freeman, 1980.

[2] H. W. Austin, *An assessment of mathematical implication in college students*. International Journal of Mathematical Education in Science and Technology **15** (1984) 327-338.

[3] M. Ayalon, and R. Even, *Deductive reasoning: In the eye of the beholder*. Educational Studies in Mathematics **69** (2008) 235-247.

[4] M. Ayalon, and Even, R. Views of mathematical educators on the role of mathematics learning in the development of deductive reasoning. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rotano, and A. Sepulveda (Eds.), Proceedings of the Joint Meeting of PME 32 and PME-NA XXX, **Vol. 2** (pp. 113–120). Mexico: Cinvestav-UMSHN 2008.

[5] S. Baker, Proofs and logic: An examination of mathematics bridge course proofs. Unpublished master's thesis. Cookeville, TN: Tennessee Technological University 2001.

[6] F. Beeftink, W. van Eerde, and C. Rutte, *The effect of interruptions and breaks on insight and impasses: Do you need a break right now?* Creativity Research Journal **20** (2008) 358-364.

[7] K. Duncker, On problem solving. Washington, D.C.: American Psychological Association 1945.

[8] B. Edwards, and M. Ward, *Surprises from mathematics education research: Student (mis)use of mathematical definitions*, MAA Monthly **111** (2004) 411-424.

[9] S. S. Epp, *The role of logic in teaching proof*, MAA Monthly **110** (2003) 886-899.

[10] V. Freiman, and B. Sriraman, *Does mathematics gifted education need a working philosophy of creativity*? Mediterranean Journal for Research in Mathematics Education **6** (2007) 23-46.

[11] B. Garii, That "Aha" experience: Meta-cognition and student understanding of learning and knowledge. Retrieved May 15, 2012, from `http://eric.ed.gov/PDFS/ED464127.pdf`

[12] J. Hadamard, The mathematician's mind. Princeton: Princeton University Press 1945

[13] G. Hanna, and M. de Villiers, *Proof and proving in mathematics education* (Discussion document), ICMI Study 19:, ZDM-The International Journal of Mathematics Education **40** (2) (2008) 329-336.

[14] P. Liljedahl, The AHA! experience: Mathematical contents, pedagogical implications (Doctoral Dissertation), Vancouver: Simon Frasier, University 2004.

[15] J. Lithner, *Mathematical reasoning in calculus textbook exercises*, The Journal of Mathematical Behavior **23** (2004) 405-427.

[16] E. Mann, Mathematical creativity and school mathematics: Indicators of mathematical creativity in middle school students, (Doctoral Dissertation), University of Connecticut, Storrs 2005.

[17] G. Miller, *The magical number seven, plus or minus two: Some limits on our capacity for processing information.* Psychological Review **63** (2) (1956) 343-355.

[18] R. Moore, *Making the transition to formal proof*, Educational studies in mathematics **27** (1994) 249-266.

[19] Rips, L. J. The psychology of proof: Deductive reasoning in human thinking. Cambridge, MA: The MIT Press 1994.

[20] M. Savic, Where is the logic in student-constructed proofs? In S. Brown, K. Marrongelle, and M. Oehrtman (Eds.), Proceedings of the 14th Conference on Research in Undergraduate Mathematics Education, (pp. 445-456). Portland 2011 OR: SIGMAA/RUME. Avaliable Online.

[21] A. H. Schoenfeld, *Measures of problem-solving performance and of problem-solving instruction*, Journal for Research in Mathematics Education **13** (1982) 31-49.

[22] A. Selden, and J. Selden, The role of logic in the validation of mathematical proofs, Tennessee Technological University Technical Report No. 1999-1, Cookeville, TN: Tennessee Technological University, 1999. Retrieved from `http://www.tntech.edu/techreports/techreports.html.`

[23] A. Selden, K. McKee, and J. Selden, *Affect, behavioural schemas and the proving process*, International Journal of Mathematical Education in Science and Technology **41** (2) (2010) 199-215.

[24] J. Selden, and A. Selden, Teaching proof by coordinating aspects of proofs with students' abilities. In D. A. Stylianou, M. L. Blanton, and E. J. Knuth (Eds.), eaching and learning proof across the grades, (pp. 339-354), New York, NY: Rutledge 2009.

[25] U. N. Sio, and T. C. Ormerod, *Does incubation enhance problem solving*? A meta-analytic review, Psychological Bulletin **35** (2009) 94-120.

[26] D. Solow, How to read and do proofs, Wiley, New York 1982.

[27] B. Sriraman, *The characteristics of mathematical creativity*, The Mathematics Educator **14** (2004) 19-34.

[28] D. J. Velleman, How to prove it. New York, Cambridge University Press 1994.

[29] G. Wallas, The art of thought. New York, Harcourt Brace 1926.

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

[31] K. Weber, *How mathematicians determine if an argument is a valid proof.* Journal for Research in Mathematics Education **30** (2008) 431-459.