When students prove a theorem without explicitly using a necessary condition: digging into a subtle problem from practice

Title: When students prove a theorem without explicitly using a necessary condition: Digging into a subtle problem from practice

Corresponding Author:

Kathleen Melhuish
melhuish@txstate.edu
Department of Mathematics
Texas State University
601 University Drive
San Marcos, TX 78666, USA
Phone: 512-245-8018 Fax: 512-245-3425

Additional Authors:

Sean Larsen
Portland State University
Fariborz Maseeh Department of Mathematics and Statistics
Portland State University
724 SW Harrison Street
Portland, OR 97201, USA

Samuel Cook
Boston University
Wheelock College of Education and Human Development
2 Silber Way
Boston, MA 02215, USA

Abstract:

Over the years, we have noticed our students constructing proofs that commutativity is preserved by isomorphism that do not explicitly use the fact that the isomorphism is surjective. These proofs are typically correct otherwise. However, such proofs are invalid because they would prove the false claim that commutativity is preserved by any homomorphism. This observation from practice raises researchable questions: How common is this phenomenon? What is the nature of this phenomenon and can we explain why students produce this type of argument? In this paper, we report a small-scale two-part survey study and a preliminary interview study designed to begin exploring these questions. Our results suggest that this phenomenon is likely quite common goes beyond a simple omission of a proof detail. Drawing on the research literature and our follow-up interviews, we propose potential explanations for this phenomenon. Finally, we discuss two different ways to think about supporting students who make this error, one that focuses on refining the students' proofs and one that involves encouraging students to use the conclusion of a statement to structure a proof.

Keywords: proof, proof frameworks, group theory, abstract algebra

When Students Prove a Theorem Without Explicitly Using A Necessary Condition:

Digging Into A Subtle Problem From Practice

In advanced mathematics courses students are asked to prove numerous theorems that deal with whether properties are preserved under certain kinds of mappings. This class of theorems is particularly important in abstract algebra and typically involves the preservation of algebraic properties by homomorphisms or isomorphisms. The following theorem is a fundamental example and its proof serves as a standard exercise in introductory group theory courses.

> If $\varphi: G \to H$ is an isomorphism and G is an abelian (commutative) group, then H is abelian.

In Figure 1, we see a correct proof of this theorem. Notice that the proof begins with the introduction of two arbitrary elements of *H* and concludes when the student establishes that these two elements commute.

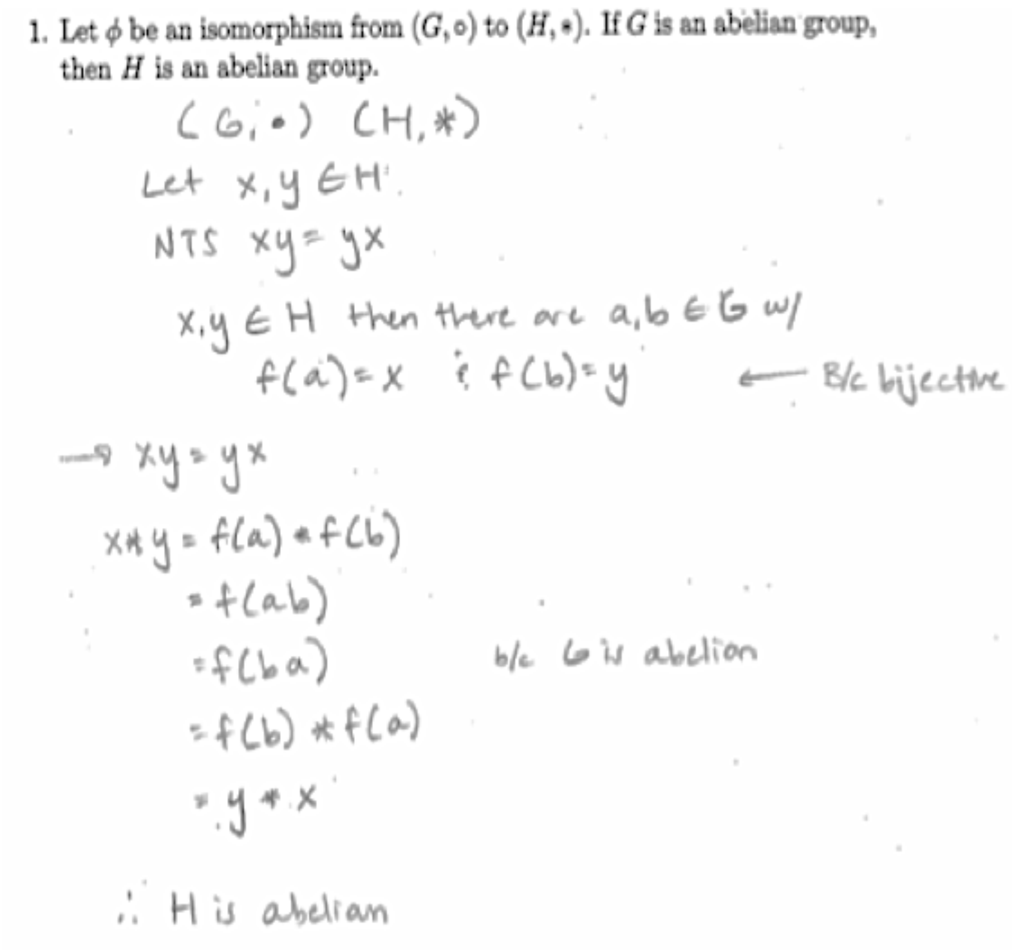
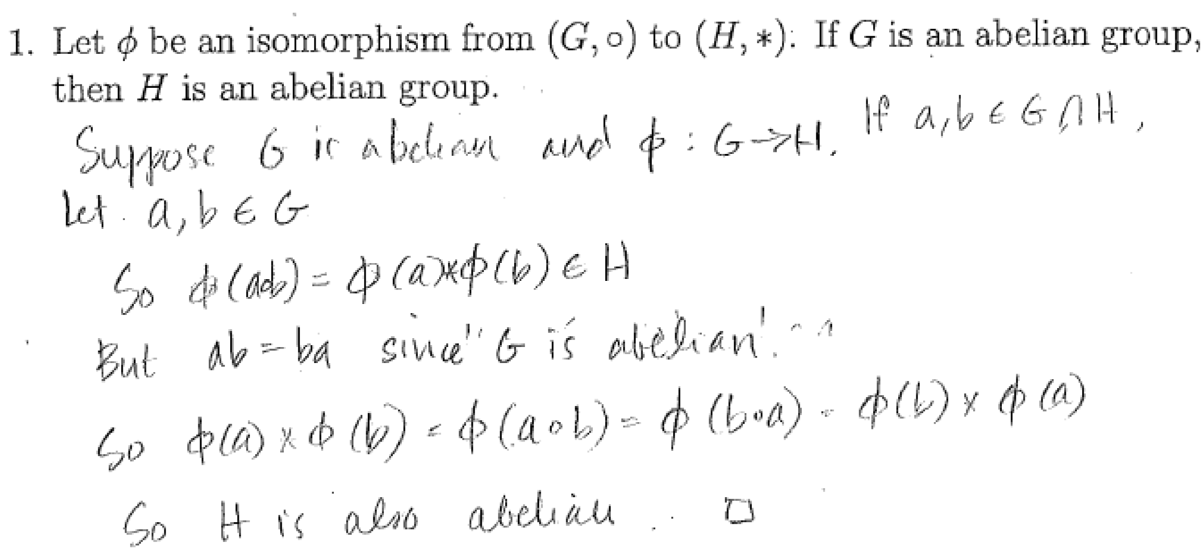


Figure 1. A valid student proof.

Over the years, while teaching undergraduate (and graduate) abstract algebra, we have observed that students commonly begin by introducing two arbitrary elements of *G* and then, using the fact that these two elements commute, show that the images of these two elements commute. Such a proof is invalid or, more precisely, incomplete without a statement that this accounts for all pairs of elements in *H* because the mapping is surjective. In Figure 2, we see such an (invalid) proof.


1. Let ϕ be an isomorphism from $(G, \circ)$ to $(H, *)$: If G is an abelian group, then H is an abelian group.

Suppose G is abelian and $\phi : G \to H$. If $a, b \in G \cap H$,
let $a, b \in G$
So $\phi(a \circ b) = \phi(a) * \phi(b) \in H$
But $ab = ba$ since G is abelian.
So $\phi(a) * \phi(b) = \phi(a \circ b) = \phi(b \circ a) = \phi(b) * \phi(a)$
So H is also abelian. $\square$

Figure 2. An invalid student proof

This observation from instructional practice was the impetus for the explanatory mixed methods study (Creswell, 2014) reported here. First, we wanted to check our perceptions as instructors by investigating and documenting how common the production of such invalid proofs actually is. Our second goal was to better understand the nature of this kind of proof writing error. It has the peculiar property that there are no errors in the proof itself. Instead the problem is that what is established is not exactly the conclusion. Or rather, it is equivalent to the conclusion for reasons that the proof writer has not stated. The proof effectively establishes that any two elements of the image of the mapping will commute. This is equivalent to the conclusion because the mapping is surjective and so the image of the mapping is the codomain *H*. In fact, some students have responded to our critiques by saying exactly this. This raises the question of whether the students were aware of this and simply neglected to mention it in their proof or were unaware and only made the connection after their proof was challenged. As we will elaborate below, the answer to this question can be helpful when considering how best to support students who produce these kinds of proofs. With this in mind, the first part of our investigation was a two-part survey study designed to both learn how common this phenomenon is and to determine the likelihood that it goes beyond a simple omission of an explicit mention of surjectivity.

A significant difference between the valid proof in Figure 1 and the invalid proof in Figure 2 is that the valid proof begins with a pair of arbitrary elements of *H* while the invalid proof begins with a pair of arbitrary elements of *G*. The second phase of our study consisted of a dive into the proof research literature followed by a handful of follow-up interviews conducted in order to help us generate potential explanations for why students might start with elements from *G* rather than from *H*. The structure of the paper reflects the chronology of our investigation. First, we describe the survey phase of our work and share the results. Then we discuss some existing research and theoretical constructs that we found to be helpful in framing this phenomenon. After introducing these ideas, we describe our exploratory interviews and discuss some conjectured explanations of why students might generate proofs that start with elements of *G*. We conclude by considering two different approaches to supporting students who do so – one that focuses on recognizing that the argument is incomplete and then refining it to produce a valid proof and another that focuses on promoting the idea of attending to the conclusion when structuring a proof.

Phase 1: Quantitative Survey Study

The first phase of our study was designed to address two questions. First, we wanted to determine how common the proof structure (starting with elements in *G*) illustrated in Figure 2 is among abstract algebra students. Second, we wanted to determine how likely it was that a student producing this kind of proof had in mind the role of surjectivity and simply neglected to include this information in the proof. We conjectured that if students were not aware of the importance of surjectivity, they would produce the same proof for the (false) statement that a one-to-one homomorphism preserves commutativity whereas if they were aware of its importance, they would instead argue that the statement was not true (and perhaps provide a counterexample). With this in mind, we designed two versions of the survey each including a different statement about the preservation of commutivity – one the true statement about isomorphisms and the other the false statement about one-to-one homomorphisms:

> *Survey Version 1*: Let φ be an isomorphism from (G, o) to (H,*). If G is an abelian group, then H is an abelian group.
>
> *Survey Version 2*: Let φ be a 1-1 homomorphism from (G, o) to (H,*). If G is an abelian group, then H is an abelian group.

Thirty-two students completed Survey A and thirty-three students completed Survey B. Students were asked to determine whether the statement was true or false and to provide a proof to support their claim. Definitions of homomorphism, isomorphism, 1-1, onto, and abelian were provided on the survey to ensure that student responses would not be constrained by their ability to recall a definition.

The survey was administered to students in three group theory classes at a public, urban university in the United States. Two of the classes were sections of an introductory group theory course meant as a first introduction to the concepts for mathematics majors (typically taken in the third year of study). The third class was a cross-listed course taken by advanced undergraduates or beginning graduate students who have typically already taken an introductory group theory course. The survey was administered at the end of the introductory course and at the beginning of the upper-level course. Most of the participants were undergraduate mathematics majors, but some students in the cross-listed course were beginning graduate studies focused on mathematics or mathematics education.

Survey Data Analysis

The first phase of analysis was to document the truth-values students assigned to the prompts. The second phase involved coding each proof attempt as either starting in the domain *G* (structuring based on the hypothesis), the codomain *H* (structuring based on the conclusion), or neither. A proof was coded as "starting in *G*" if the student began by introducing a pair of elements in *G* and coded as "starting in *H*" if the student began by introducing a pair of elements in *H*. The purpose of this second kind of coding was to determine how strongly starting a proof with elements of *G* was associated with the production of an invalid (incomplete) proof. (See Table 1.)

When students presented a deductive argument, it was coded as *invalid* or *valid*. Note that any deductive argument for the Version 2 statement (that attempted to establish the truth of the statement) was necessarily invalid due to the statement being false. In order to assess whether a proof of the Version 1 statement was valid, two considerations were used. The first consideration was whether the proof was structured in a way that could establish the truth of the result. Thus, a proof that began with two arbitrary elements of *H* and attempted to established that these commuted was considered to have a valid structure. Additionally, a proof that started with elements of *G*, established that images of these elements commuted, and then cited surjectivity to establish the conclusion was also considered to have a valid structure. The second consideration was at a line-by-line level. We evaluated each proof to determine if it each line followed logically from the previous line. However, consistent with Inglis, Mejia-Ramos, Weber, and Alcock (2013) we recognize that the level of depth and detail needed for a proof to be considered valid varies depending on any number of factors including the context of the proof and the mathematical background of the reader. A code of "invalid" was only assigned to an argument featuring an appropriate structure in cases where major flaws were identified such the use of false statements or lack of deductive argument. Consider Figure 1. While the student could have been explicit citing the homomorphism property as a warrant for two of their steps, the proof was considered valid in our analysis.

A second coder, a mathematics education Ph.D. student, coded a random subset of fifteen surveys. Agreement over validity was 100%, and agreement on codes for *G*-first or *H*-first frameworks was 93.3%. (The disagreement involved an argument that was extremely incomplete and included statements that could either be interpreted as evidence of starting in *G* or as evidence merely of the student writing down a translation of the statement to be proven. Through discussion, we reached an agreement and treated the statement as *G*-first.) After the surveys had been coded, a proportion hypothesis test was conducted on the deductive arguments to explore if the *G*-first approach occurred at a rate greater than chance. We then explored the hypothesis that starting in *H* would lead to more valid proofs by comparing the frequency of students presenting valid proofs starting in *G* with the frequency of students presenting valid proofs starting in *H*.

Survey Results

We present the results of our survey study in Table 1 (Version 1) and Table 2 (Version 2). For each of the three classes, we report the number of students who responded, the number of students who assigned the correct truth-value to the prompt, the number of students who provided a valid proof, and the number of students who provided arguments coded as starting in *G* (i.e. structuring based on the hypothesis). Class A and B were introductory level, and class C was the advanced class.

Table 1. Version 1 (True Statement)

Class	Number of Students	Correct Truth Value (T)	Valid Proof	G-first framework
A	8	8	1	3
B	8	7	4	3

C	16	13	4	10
Total	32	28	9	16

Table 2. Version 2 (False Statement)

Class	Number of Students	Correct Truth Value (F)	Valid Proof	G-first framework
A	8	0	0	6
B	8	1	0	4
C	17	2	1	13
Total	33	3	1	23

As seen in Table 1 and Table 2, students tended to assign a value of *true* to both prompts. In fact, no student incorrectly assigned a value of *false* to the Version 1 prompt (twenty-eight students labeled the prompt *true,* while the remaining students failed to assign a truth value at all). Only three of the students responding to Version B correctly identified the prompt as false.

Of those, thirty-nine arguments were coded as starting in *G* and seventeen were coded as starting in *H*. Students presented arguments beginning in *G* at a rate that, when compared to chance, was significantly greater (p=0.0016) and has a moderate effect size (h=.4936). Assuming a type I error of 5% (α=0.5), this sample size and effect size projects a power of 0.9798.

In response to the false statement, thirty-one students attempted a deductive argument to prove the statement true. Clearly all of the deductive arguments for the false statement were invalid. Of these students, twenty-three used a *G*-first proof framework. It is important to note that these proofs were actually identical to the incomplete proofs we saw for the true statement. Consider the proof in Figure 3.

1. Let ϕ be a 1-1 homomorphism from $(G, \circ)$ to $(H, *)$. If G is an abelian group, then H is an abelian group.

pf: Let ϕ be a 1-to-1 homomorphism from $(G, \circ)$ to $(H, *)$.

Then $\forall a, b \in G$, $\phi(a \circ b) \overset{①}{=} \phi(a) * \phi(b)$.

Suppose that G is an abelian group. Then

$$\phi(a) * \phi(b) = \phi(a \circ b) = \phi(b \circ a) = \phi(b) * \phi(a),$$

(↑ eqn ①, ↑ G abelian, ↑ homomorphism)

so $\phi(a) * \phi(b) = \phi(b) * \phi(a)$ $\forall \phi(a), \phi(b) \in H$, and $\therefore$ H is abelian. ☒

This result, especially in light of the very small number (three) of students who correctly identify this statement as false, suggests that is common for students to be unaware of the role of surjectivity in preserving commutativity under isomorphism. This suggests that the incomplete proofs of the true statement are unlikely to be the simple result of students' neglecting to make their use of surjectivity explicit. Thinking ahead to implications for practice, this finding is significant because it means that supporting students in refining these proofs will likely require helping them to realize that the result depends upon surjectivity.

Because the isomorphism prompt was a true statement, we conducted further analysis of these proof attempts. For this prompt, twenty-five students attempted deductive arguments. We compared *G*-first frameworks (structured around the hypothesis) and *H*-first frameworks (structured around the conclusion) of these arguments. Sixteen started in *G* and nine started in *H*.

Of the sixteen who started in *G*, only two students provided valid proofs. Of the nine students that started in *H*, seven provided a valid proof. A Fisher Exact Test indicates the proportion of students who provide valid proofs is significantly higher for those who utilize an *H*-first structure; this test is complimented by a large effect size and large odds ratio (p=0.0022, ϕ= 0.653, the odds ratio is 20.1). The odds ratio estimates that those who use an *H*-first framework are 20.1 times more likely to provide a valid proof. With a small sample size the odds ratio has a large variance; however, the lower bound of a 95% confidence interval for the true odds ratio is 2.1.

These results highlight that most of the students began their proof with elements of *G* and most of these produced invalid (incomplete) proofs. At the same time, those few students who did begin with elements from *H* were very likely to produce valid proofs. Note that in almost all deductive proof attempts (both versions), students were successful in using the commutativity of *G* in concert with the homomorphism property to show the image elements commutated. This suggests that almost all of the students had sufficient understanding of the definitions of homomorphism and commutativity to successfully produce a proof.

In summary, the production of proofs like the invalid proof in Figure 2 is likely quite common among abstract algebra students and most such proofs are incomplete not because of a simple omission but rather because the students are unaware of the role of surjectivity in preserving commutativity. Further, the choice of whether to begin with elements in *G* or with elements in *H* was highly predictive of student success in constructing valid proofs. For this reason, it is worthwhile to attempt to understand why students might be so likely to attempt an approach that is unlikely to lead to a valid proof. To address this question, we first discuss some of the proof literature and introduce the notion of a proof framework to provide a lens to more precisely characterize the phenomenon we observed in our survey data. Then we describe our exploratory interview study and share some possible explanations for why so many students approached the proof by starting with elements from the domain group *G*.

Literature Background on Student Proof Construction

In this section, we provide an overview of literature about students' proof construction to situate the phenomenon we documented in the survey study. In general, the research has illustrated that producing valid proofs is a complex and challenging endeavor for students (Stylianides, Stylianides, & Weber, 2017). We situate our research in one key aspect of proof construction: producing an argument that appropriately aligns with the statement to be proven. In exploring how undergraduates and experts approach proof tasks, Alcock and Weber (2004) illustrated two ways of evaluating and proving/disproving a statement: through syntactic or semantic exploration. They found that many students operate syntactically by immediately attempting a formal proof via setting up assumptions, unpacking definitions for the statement, and then making a series of implications without any additional exploration of the statement to be proven. In extreme examples, students may be swayed by an argument's validity simply because it appears to mimic this syntactic structure (e.g., Harel & Sowder, 1998). Selden and Selden (2003) further found that undergraduates focused on such surface features of purported proofs with little attention to whether the general structure was appropriate to prove the given statement. These results reflect that students may not be meaningfully exploring statements to be proven or meaningfully analyzing proofs to assure alignment with the statements.

Selden and Selden (1987) specifically classified the error of students producing proofs that are valid arguments, but insufficient to prove the intended statement as cases of *weakening the theorem.* In these cases, a student proves a result that is weaker than the conclusion or uses assumptions stronger than those in the hypothesis. Such is the case in our example when students produce arguments about the image of an isomorphism rather than the intended co-domain. The activity of weakening the theorem can be quite productive, and mathematicians weaken and strengthen theorems often in their practice (Lakatos, 1976). However, this is rarely asked of students (Jahnke, 2008), and thus it is unlikely that they consciously analyze proofs for potential theorem weakening.

To further classify the nature of the *H*-first, and *G*-first approaches to our prompt, we turn to Selden and Selden's (1995, 2015) notion of *proof framework*. A proof framework is the "representation of the 'top-level' logical structure of a proof" (Selden & Selden, 1995 p. 129). Selden and Selden documented that students struggle to move between a given statement and the appropriate proof framework. If we return to our valid statement about isomorphism preserving commutativity, an appropriate proof framework would be as follows:

Suppose G is abelian and $\varphi: G \to H$ is an isomorphism.

Let a, b be elements of H.

[ ]

Then $ab = ba$.

Therefore, H is abelian.

In contrast, the proof framework frequently employed by our participants looked like:

Suppose G is abelian and $\varphi: G \to H$ is an isomorphism.

Let a, b be elements of H.

[ ]

Then $ab = ba$.

Therefore, H is abelian.

The *first-level* proof framework for both approaches is the same: assume that the hypothesis of the statement is true in the first line and conclude that the conclusion is true in the last line. (Note, that is it not unusual for one or both of these lines to be left implicit in a proof as we see in the proofs shown in Figure 1 and Figure 2.) Where the two structures diverge is at the *second-level* prof framework. The second-level framework (of an implication) results from unpacking the statement of the conclusion. In this case, the conclusion is "H is abelian," a statement that should be unpacked as, "Let a, b be elements of H… Then $ab = ba$." Using this terminology, we can precisely describe the error in the proof shown in Figure 2 by saying that it utilizes an inappropriate second-level proof framework: "Let a, b, be elements of G… then $\varphi(a)\ \varphi(b) = \varphi(a)\ \varphi(b)$." Specifically, the first step in the argument, "Let a, b in G" is not part of what it means for H to be abelian. Rather, it comes from unpacking the hypothesis. The last part of the second-level frame, "$\varphi(a)\ \varphi(b) = \varphi(a)\ \varphi(b)$," is also inappropriate as it refers to the image of the map and not the codomain group H. In this situation, a student can produce a proof that appears to be valid (and is a valid proof of a different statement), but is fundamentally misaligned with the statement to be proven. Now that we have leveraged proof frameworks to identify what our phenomena is a case of, we turn to follow-up interviews to provide explanatory power for the prevalence of the students' proof structuring decision.

Phase 2: Interview Study

We conducted follow-up interviews with nine of the students submitting surveys with the goal of constructing explanations for the various response types using a semi-structured interview approach. During the interviews, students were shown their responses and asked to recall and explain their thinking as when they approached evaluating and proving the statement. The interview included explicit prompts about why a student began in G or H. We divide our results into two sections. First, we briefly share two cases of students with a valid approach to version A and version B, respectively. These cases were outliers but may provide insight into what differentiated them from their peers. We then present two in-depth cases of students engaging in the more typical *G*-first approach. We selected these students because they were able to fully explicate their rationale for beginning with a *G*-first approach. After providing a narrative of each students' approach and explanation for it, we introduce the *met-before* construct (McGowen & Tall, 2010) to articulate two plausible explanations for their proof structuring decisions.

Valid Approach to Version 1 (The True Statement)

First, we consider Beth, an undergraduate student, who provided the valid proof to the version A prompt found in Figure 1. Her explanation for her method of structuring the proof is as follows:

> Beth: Suppose *G* is an abelian group. In particular, that's what it means to be abelian for all elements of *G.* Each element will commute so I think I reference it from the definition you gave. Since phi is an isomorphism from *G* to *H* for all *a*, *b* in *G*. This is true, so it preserves the operation. Then I let two elements be in *H* because we want to prove that *H* is commutative for all elements in *H.*

In this description Beth hints at her first-level framework when she says, "Suppose G is an abelian group…" We also see an explicit statement indicating a focus on establishing the truth of the conclusion: "…then I let two elements be in *H* because we want to prove that *H* is commutative for all elements in *H*." Beth's approach instantiates using a second-level proof framework by attending to the conclusion.

Valid Approach to Version 2 (The False Statement)

Because her version of the survey featured the true statement, Beth's valid response was in the form of a deductive proof. In contrast, a valid approach to the false statement (Version 2), involves generating a counterexample to the claim. Because of the dearth of correct responses providing a counterexample, we share pilot data from a prior term with Michael, a graduate student who had completed the advanced course.

> Michael: What I've done in here, is that I have... let's call this $\mathbf{Z}_4$ (draws function diagram similar to that in Figure 3), and I have $\mathbf{D}_8$ here, and I have a little $\mathbf{Z}_4$ sitting inside it. Let's map that to that. This part is abelian. This part is abelian. But the whole thing isn't abelian.
>
> Interviewer: Right, so is this what you were thinking when you were coming up with your argument?
>
> Michael: Yeah, I think so.
>
> Interviewer: This sort of diagram in your head?
>
> Michael: Yeah, that I could send my abelian group into a group that isn't abelian but that has an abelian subgroup.

Michael took a semantic approach (Weber & Alcock, 2004) to evaluating this statement. By that, we mean that he left the formal system to explore the statement via examples and diagrams. The counterexamples to this prompt are numerous and could be as simple as mapping the trivial group to the identity element of any non-abelian group. The key is to address the consequences of the function not being surjective. Michael used the dihedral group of order 8, a group he knew to be non-abelian with an abelian subgroup isomorphic to $\mathbf{Z}_4$. To generate a counterexample, he considered the obvious map from the abelian group $\mathbf{Z}_4$ into $\mathbf{D}_8$ whose image was the cyclic subgroup of order 4. The rarity of such semantic arguments in our data may reflect novice students' preference for producing proofs syntactically without exploring the given statement (Weber & Alcock, 2004).

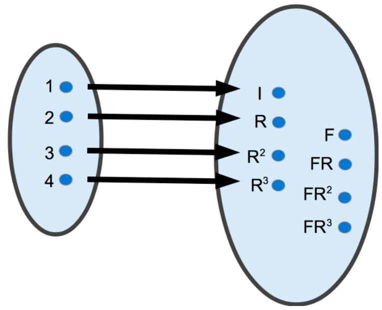


Figure 3. Michael's Diagram illustrating a 1-1 homomorphism from an abelian group to a non-abelian group.

Insight from Valid Approaches

While these types of valid responses were outliers in our study, we can glean some insight into what supported their success from our interviews with them. In both cases, the arguments produced aligned directly with the statement to be evaluated. Beth's explanation suggests that a focus on the conclusion of the statement may have supported her in constructing a valid proof utilizing an appropriate second-level proof framework. Michael's semantic approach to generating a counterexample suggests that he was able to coordinate a sufficiently rich example space (Sinclair, Watson, Zazkis, & Mason, 2011) with his understanding of the relevant concepts (homomorphism, abelian) to consider the veracity of the statement. However, it seems likely that the knowledge Michael and Beth drew upon was also available to the unsuccessful students. The knowledge needed to unpack the statement that G is abelian (which most of the unsuccessful students did successfully) is the same as needed to unpack the statement that H is abelian. And the examples Michael used are not at all esoteric and were likely familiar to most of the students. Thus, it seems likely that the other students *could* have used the approaches of Michael and Beth if they had reason to attempt them. Put differently, it seems unlikely that the students chose invalid approaches because they lacked the skills or knowledge needed to execute a valid approach. Rather, the strategy employed seemed to be the correct way to approach the task.

What we hoped to uncover in our interviews with students who provided invalid arguments, was insight into the reasoning that supported their use of hypothesis-driven second-level proof frameworks. We present two cases that reflect differing explanations for why students might begin their proofs in this manner. After presenting the cases, we use McGowen and Tall's (2010) construct of met-before to frame the students' thinking. We argue that there are two mutually reinforcing phenomena that may account for the kind of second-level proof framework that was commonly observed in our study - one tied to the structuring of proofs of if-then statements in general, and one tied more specifically to the function concept.

Case 1: Rebecca

Rebecca was a student who completed the survey at the end of the introductory group theory course. The proof she provided on the survey is presented in Figure 4. Rebecca's proof began by unpacking the assumptions including phi being a homomorphism and 1-1, and *G* being abelian. Her second-level proof framework was structured by the hypothesis where she leveraged elements of *G* commuting to argue that their images also commute, arriving (invalidly) at the conclusion that *H* was abelian.

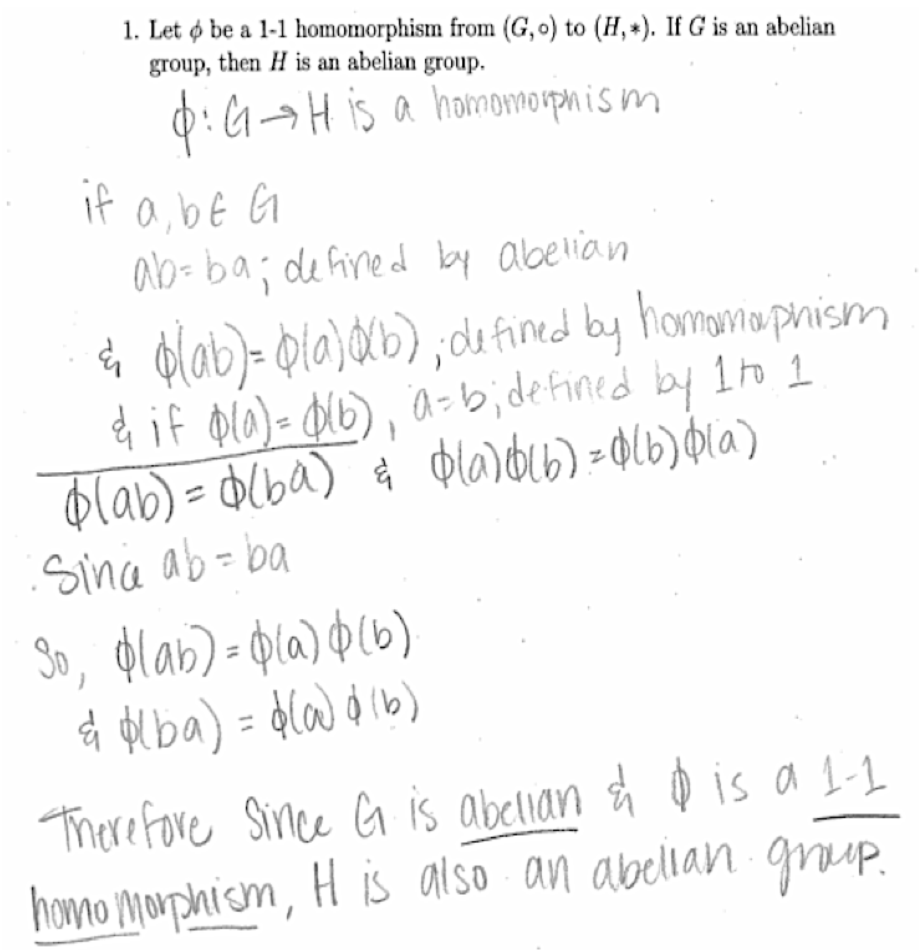


Figure 4. Rebecca's proof for Version B (False Statement)

When prompted to talk through her process she began:

Rebecca: Okay, I started by defining what the homomorphism is, and that's with just the mapping here. That's the definition. And I kind of went through to define what I know. So, if it's abelian, *ab*=*ba*. And so, I took that, and it's one to one so with the homomorphism you have an *e*. We're trying to prove that if *G* is abelian, then *H* is also abelian.

Rebecca was then asked, "So, why did you decide to start in *G*?" She responded by saying that:

G is the group that we are mapping from, and I guess with the 1-1 the way that my mind always is that you take your elements from that and move to the other group. So, the mapping is from *G,* so I'm taking my elements from *G*.

Her response reflected a domain-first approach when dealing with the homomorphism function. She noted that she always approaches 1-1 in this manner as well. She was asked to consider starting in *H* instead.

> Interviewer: So, what if I said for my very first step, "If *c* and *d* are in *H*"?
>
> Rebecca: Then you would still follow the same equation, but you would be working backward instead.

A few moments later, we have evidence regarding what she meant by working backward when she asserts that in fact, she would need to use the inverse of the homomorphism if she began in *H*.

> Interviewer: So, I'm going to push you on that one a little bit. If we can go either way, why do we set it up so we are just going from *G* to *H*?
>
> Rebecca: [Be]cause we are [pause]- We are taking our elements from a group and mapping them to a separate group and if this states that is 1-1 homomorphism from *G* to *H* with our function phi going [pause]- I guess it wouldn't be the exact same one it would be an inverse then.
>
> Interviewer: What do you mean by an inverse?
>
> Rebecca: It's going to be doing - it's undoing the function that's going. We are doing one function that is taking these into these and this; we are basically undoing this function to bring it back to here.

We see further evidence of a domain-first directionality in a description later in the interview of why she felt that commutativity of *G* was preserved in *H*:

> Rebecca: If all the elements in here [*G*] can be moved around in that respect of $ab = ba$ then functions going- your ϕ will preserve those elements - the characteristics of that element so that over here you're going to get - this here $\phi(a)\phi(b) = \phi(b)\phi(a)$.

From this description, Rebecca appears to be imagining the commutativity in *G* being transmitted to *H* by the homomorphism.

What she never seemed to do was consider the validity of the conclusion itself. In fact, when asked explicitly to address whether all of *H* must be abelian, Rebecca was able to use her knowledge of onto to reassess the truth of the original statement.

> Interviewer: What is our goal? We are trying to prove what?
>
> Rebecca: To prove that *H* here is abelian if *G* is abelian.
>
> Interviewer: So, okay. I'm seeing that you have - you made an argument that everywhere you map to swapping them - the question is - *does that mean all of H is abelian*?
>
> Rebecca: I guess it - Since it doesn't say that it is onto there could be elements inside *H* that are not part of that phi. Part of the phi function mapping. So, then those elements would not be necessarily abelian.

Even after this realization, Rebecca was unsure how she would approach the proof.

> I'm not sure. [extended pause]. I don't know where I would have the other elements of *H*. Like how I would show in here that *H* also has elements. That would true that all of the ones from *G* to *H* would indeed be abelian but the other elements of *H* may not. I wouldn't know how to put that into respect with making this proof. Cause I'm using just my definitions.

Rebecca remained unsure of how to work her new insight into a proof. This may reflect a difference between her and Michael's proof production approach. Rebecca generally remained syntactically focused when proving. Approaching the proof syntactically from the perspective that functions (especially 1-1) have directionality (they go from the domain to the codomain), it made sense to Rebecca to start with elements from *G* and then apply definitions to illustrate the commutativity moving from *G* to *H*.

Case 2: Jessica

Jessica was another undergraduate student from the introductory proof course. The proof she provided on the survey is presented in Figure 5. Jessica's proof was similar in structure to Rebecca's proof. Like Rebecca, Jessica began her discussion of her proof by setting up definitions and interpreting the statement. Her first-level proof framework begins with the hypothesis and ends with the conclusion. Her second-level proof framework is structured by a statement of what she wants to show if *a*, and *b* are in *G,* then the product of their images commutes. We acknowledge that the proof contains some additional errors including misplaced and insufficient warrants. For the scope of this paper, we focus the exclusively on her proof framework decisions.

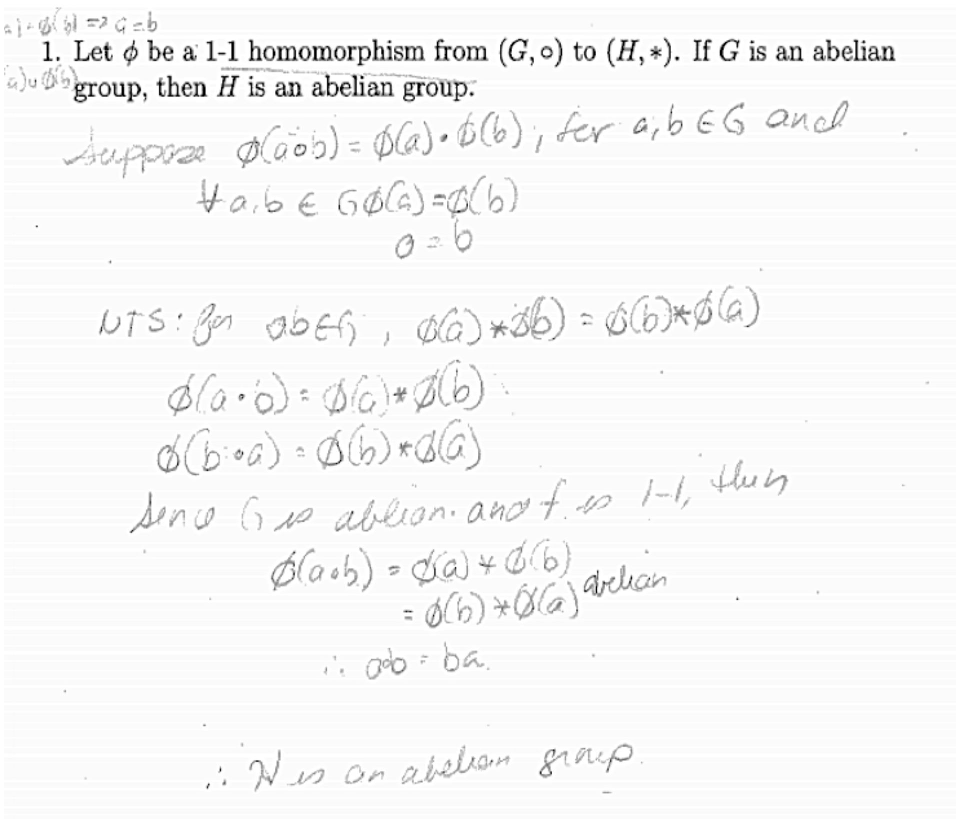


Figure 5. Jessica's proof for Version B (The False Statement)

Jessica explained how she addressed the prompt:

> Okay, so because this is the assumption – that is a 1-1 homomorphism that's actually a given, the assumption that *G* is abelian, so I need to prove that *H* is an abelian group.

As she reflected on her proving process, she pointed to the set-up of her proof where she defined abelian and homomorphism. As in the case of Rebecca, Jessica then introduced arbitrary elements in G, stating, "So I need to show that whenever elements a and b – when it's mapped into H that, that's abelian." The interviewer asked why she began with elements in G.

> Jessica: I think it was a bit robotic because I was so new at proofs and I had worked with a private tutor, and I had seen my instructor in office hours. It sort of began robotic, and I had to adopt a robotic method to understand that when you are given a proof of one of the first things you wanted to do is take your assumption and introduce elements specifically for group theory. So, that's where that came in.
>
> Interviewer: Okay, so you always begin with your assumptions and work down...
>
> Jessica: I always begin with my assumptions and work down to my conclusions. In group theory, I also introduce elements in my initial group.

Jessica's response indicated a desire to begin with the assumptions or givens for a prompt. She then chose her arbitrary elements from the *if* side of the *if-then* statement. Her statement that she introduced elements in the "initial group" could be interpreted to mean she introduced elements from the domain group (making her reasoning consistent with Rebecca's). However, we argue that it is much more plausible that by "initial group" she means the group that appears in the hypothesis. We base this argument on two things. First, Jessica's language is focused on assumptions and conclusions with no mention of functions at all. Second, she refers to group theory in a general way that refers to statements about groups without any restriction to statements that involve functions. Almost any if-then statement in group theory is likely to mention a group in the assumption, but only a subset of statements is likely to involve a map between two groups.

The interviewer followed up by asking if that is how she always attacked a proof. She responded:

> I always start with my assumption, and I used different mathematical logics to get my [conclusion]. Initially, I tr[ied] to introduce my assumption and work sometimes from the conclusion and prove it and I learned quickly that was wrong, so I know from – I need to manipulate and apply theorems to get the part that needs to be proved.

Jessica reported that she made her structuring choice because she had learned always to begin with the assumptions and work to the conclusion. As in Rebecca's case, Jessica appeared to work syntactically, leveraging her knowledge of creating deductive arguments for if-then statements.

Met-Before Explanations for Hypothesis-Driven Proof Frameworks

McGowen and Tall (2010) define a *met-before* to be a "mental structure that we have now as a result of experiences we have met-before." (p. 171). Met-befores can be helpful when students learn new mathematics that is consistent with mathematics they have learned previously. For example, McGowen and Tall present the met-before: 2+2=4. This knowledge can help in situations such as in the context of larger numbers (e.g., 200 + 200= 400) or combining like terms in algebra (e.g., $2x+2x = 4x$). However, met-befores can also be problematic. McGowen and Tall illustrate problematic met-befores with a number of examples related to negative numbers. In one case, students were presented with a number line containing y to the left of 0. The students determined that $2y$ was larger than y even though y was a negative number. In this illustration, the students' met-before was the idea that the product of two numbers is bigger than either factor, which is true for whole numbers but not true more generally.

A met-before is particularly robust type of prior mathematical experience. Met-befores can serve as epistemological obstacles that overwhelm the influence of more recent experiences. The students in this study have likely seen proofs structured by their conclusion as well as proofs starting in the co-domain, but it is unlikely that such recent experiences would have coalesced into robust met-before at the time of the study. We argue that, despite likely supportive recent experiences, Rebecca and Jessica were each hindered by a problematic met-before when they formulated their proofs.

Rebecca's Problematic Met-Before: Function Directionality

Rebecca began her proof production with elements from *G* and then mapped them to *H*. Her reasoning for this decision was based explicitly on her past experience with mappings. She stated that "the way that [her] mind always is" was to start with elements from the [domain] group and map to elements in the [codomain] group.

The mathematics education community has extensively studied students' conceptions and activity dealing with complexities of function mappings (Dubinsky & McDonald, 2002; Oehrtman, Carlson, & Thompson, 2008). In Rebecca's case, she appears to conceive of functions as one directional where elements begin in the domain and then are mapped to the range. This is in contrast to an understanding of a function as a relationship between a pair of elements where one could begin with something in the range and find elements that mapped to those outputs. Oehtrman, et al. noted the difficulty in dealing with the preimage of a function (emphasizing the need for a holistic treatment of a function, a *process* conception.) In, Rebecca's case, when prompted to address elements from the co-domain, she introduced an inverse function in order to change directionality, rather than treating the preimage as part of the original relation.

This desire to begin in the domain has been documented in advanced mathematics students in other contexts, notably by Swinyard and Larsen (2012) in the context of limits. In their teaching experiment which focused on students reinventing the formal definition of limit, moving from an *x*-first to a *y*-first perspective was a major transition. The students they worked with had a strong preference for starting with inputs rather than outputs. These students may have been battling with the same met-before as Rebecca: function directionality determines which part of a function we attend to first: domain, then codomain.

Jessica's Problematic Met-Before: Second-level Proof Frameworks Based on Hypothesis

Jessica explained her process as linked to her standard strategy for proving if-then statements. She claimed she always begins with assumptions and arrives at the conclusion. This strategy is appropriate for first-level proof frameworks and frequently will lead to success in proving simple statements (Selden, McKee, and Selden, 2010). In the interview, Jessica was explicit about her prior experience reflecting this strategy as a met-before. Weber associated the term *strategic knowledge* for this type of decision-making. He defined strategic knowledge as, "heuristic guidelines that they can use to recall actions that are likely to be useful or to choose which action to apply among several alternatives" (p. 111). In Jessica's case, she had a strategy, to begin with assumptions that were supported by a heuristic to avoid beginning with conclusions.

Jessica's approach is not surprising in light of literature on student errors when proving. Selden, et al. (2010) found a similar error made by a student when evaluating a statement about the image of a homeomorphism. They hypothesized that idea that one should start with hypotheses (and never conclusions) was responsible for these types of errors and may have been emphasized in early proving courses such as geometry. When Jessica stated she had previously tried to "work sometimes from the conclusion and prove it and I learned quickly that was wrong" she might have been referring to the error of assuming the conclusion. The fact that a second-level proof framework comes from the conclusion of a statement could be counter-intuitive for students who have repeatedly been warned not to start with the conclusion. While there is a stark difference between *structuring* your proof based on a conclusion and *assuming* the conclusion, students may be unaware of this important distinction.

Discussion

Our study highlights several issues that are relevant to arguments about *if-then* statements. The survey results documented that the majority of students did not structure proofs based on the conclusion to be proven, suggesting that this may be a common problem for students learning to write proofs. By using a hypothesis-driven (*G*-first) second-level proof framework, the students largely produced invalid proofs (or more precisely, incomplete). In contrast, the majority of students who used a conclusion-driven second-level proof framework produced a valid proof of the given statement, suggesting that focusing on the conclusion is an important aspect of successfully proving if-then statements. Our analysis reflected that students using a conclusion-driven framework were statistically significantly more likely to produce a valid proof with a notable effect size. For this reason, we believe it is important to understand why students might choose hypothesis driven proof structures and think about how we might encourage them to choose conclusion driven proof structures. However, we also wish to emphasize that the students demonstrated many successful actions in their proof construction. While we value students having exposure to the normative ways of structuring proof, later in this discussion we also explore an alternate route of productively building from the students' *G*-first approaches.

While it has been established that students often lack strategic knowledge and knowledge of structuring proofs (Selden, et al., 2010; Weber, 2001), the issue has been explored with limited scope. Our analysis provides not just documentation of an ineffective strategic choice, but also an explanation for why students make this structuring choice. Selden, et al. reported on a single case in a topology context where a student analogously began in the domain instead of the codomain while attempting to prove a statement about a homeomorphism. They attributed this structuring decision to *behavioral schemas,* "[e]nduring mental structures that link situations to actions, in other words, habits of mind, that appear to drive many mental actions in the proving process" (p. 199). They hypothesized that working from assumptions to conclusions is a behavioral schema reinforced in courses such as geometry where this type of action often results in successful proofs.

We have expanded on this work, by not just addressing behavior, but unpacking the reasons that can account for the development of these schemas. Jessica's case aligned with Selden, et al.'s hypothesis of if-then statements being linked to hypothesis-to-conclusion structuring; however, Rebecca's structuring choices were tied to the specific content of functions rather than just the statement's if-then structure.

The construct of met-before provided us with a theoretical tool to make sense of both Jessica's content-independent and Rebecca's content-dependent proof structuring decisions. This approach contrasts to majority of proof studies that address either content-dependent or content-independent aspects of proving (Dawkins & Karunakaran, 2016). By adopting a lens that allowed for flexible exploration, we were able to make two qualitatively different plausible explanations about the reasoning that produced the same common proof-structuring issue. First, students' previous experiences involving the logical directionality (hypothesis-to-conclusions) of if-then statements became a problematic met-before. These experiences likely include instruction meant to discourage students from assuming conclusions. The logical directionality met-before may have a strong influence over how students structure their proofs. Said more simply, students' prior experiences encourage them to structure their proofs in a way that emphasizes beginning with the hypothesis and discourages them from attending to the conclusion. In our interview with Jessica, she explicitly attributed her proof structuring process to her prior experiences proving if-then statements. Second, students' prior experiences with the concepts involved in the statement may discourage them from using appropriate second-level proof frameworks. Our analysis of Rebecca's interview suggests that her met-before of function directionality encouraged her to start her proof by focusing on the domain of the function. As we noted above, other studies have documented challenges experienced by students when faced with situations where it is necessary to focus one's attention first on the codomain of a function. More generally, concept understanding and previous experience with concepts can lead to invalid proof attempts (e.g., Edwards &Ward, 2004; Moore, 1994). Our study suggests that the concepts involved in statements to be proven come with their own sets of met-befores that may impact the actions students take when evaluating and proving statements.

In this paper, we extended the literature on students' proving in two ways:

1. We provided evidence that it is common for students to construct invalid proofs (of if-then statements) with second-level proof frameworks structured by the statement of the hypothesis rather than the statement of the conclusion.
2. We provided plausible explanations for why students may or may not structure proofs in this manner.

Our analysis suggests that in general, students' prior experiences encourage them to focus on the hypothesis when constructing proofs of if-then statements. Our analysis also suggests that in the case of statements involving functions from sets mentioned in the hypothesis to sets mentioned in the conclusions, students' prior experiences with functions encourage them to structure their proofs beginning with elements from the domain.

Based on these findings, we suggest researchers interested in students' proof production attend to the ways that students' prior experiences may be supporting or constraining their efforts to produce valid properly structured proofs. For example, it may be productive to conduct design studies to develop instructional strategies to support students in learning how the syntax of a statement can inform the structuring of a proof. Such research could be seen as a quest to identify and design prior experiences that would serve as supportive met-befores as students advance as provers.

Our findings also have practical implications for teaching. For example, it may be helpful for an instructor to explicitly discuss with their students the role of conclusions in structuring proofs. Further, our findings suggest that instructors should take care that admonishments to avoid assuming conclusions do not have the undesirable side effect of discouraging students from attending to conclusions as they structure their arguments.

We can also reframe this phenomenon from an erroneous use (or lack of use) of the conclusion to a productive step in the proving process. The students in this study implicitly weakened the theorem (they proved that the image of a commutative group under an isomorphism is commutative) and produced incomplete proofs of the intended statement (that commutativity is preserved under isomorphism). There were often no errors in these arguments aside from the fact that they did not establish the conclusion. Instructors may want to explicitly attend to this situation. Proofs such as the ones in this study can lead to productive discussion about checking proofs for hidden assumptions, exploring statements to check alignment with proof, and determining alterations to a statement to move from an invalid statement to a valid statement. Such proof analysis (Lakatos, 1976) is common amongst mathematicians, but rarely asked of students (Jahnke, 2008), although research has offered existence proofs demonstrating that students can productively engage in such activity (Larsen & Zandieh, 2008).

In summation, this series of studies documented several important aspects of student proving. First, our survey analysis illustrated that students frequently used the hypothesis rather than conclusion to structure their proofs. While this approach could be patched (by appealing to surjectivity in the case of our prompt), we conjectured that students were not just overlooking a detail during their proof production. We verified this claim by illustrating that students produced the same arguments for an invalid statement without the available surjectivity patch. In this way, the lack of appeal to surjectivity was not just a detail omission, but a fundamental flaw in the alignment of the statement and produced proof. Further, from this data we established that structuring a proof based on the hypothesis of a statement was deleterious to the production of a valid argument. Using the conclusion to structure the proof much more frequently led to valid approaches. Student interviews provided potential explanation for the prevalence of hypothesis (*G*-first) second-level proof frameworks. Students prior experiences with functions may foster a domain-first approach. Students prior experiences with simple if-then statements may foster a hypothesis-first approach. Both explanations capture met-befores that were likely successful across many tasks and many settings. These diverse explanations imply that researchers may want to more fully attend to how students' experiences related to both proof and the mathematical content of statements may support or constrain their proof decisions. Further, this work was situated in one of the most ubiquitous proof tasks in abstract algebra. As such instructors may want to leverage this particular context to provide opportunity for proof analysis and opportunity to discuss the consequences of proof-structuring decisions.

References

Antonini, S., & Mariotti, M. A. (2008). Indirect proof: what is specific to this way of proving? *ZDM*, *40*(3), 401-412.

Creswell, J. W. (2014). *A concise introduction to mixed methods research.* Thousand Oaks, CA: Sage Publications.

Dawkins, P. C., & Karunakaran, S. S. (2016). Why research on proof-oriented mathematical behavior should attend to the role of particular mathematical content. *The Journal of Mathematical Behavior*, 44, 65-75.

Dubinsky, E., & McDonald, M. A. (2002). APOS: A constructivist theory of learning in undergraduate mathematics education research. In D. Holton et al. (Eds.), *The teaching and learning of mathematics at university level* (pp. 275-282). Springer Netherlands.

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

Fiallo, J., & Gutiérrez, A. (2017). Analysis of the cognitive unity or rupture between conjecture and proof when learning to prove on a grade 10 trigonometry course. *Educational Studies in Mathematics*, 1-23.

Garuti, R., Boero, P., & Lemut, E. (1998). Cognitive unity of theorems and difficulty of proof. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 345–352). Stellenbosch, South Africa: University of Stellenbosch.

Inglis, M., Mejia-Ramos, J. P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in cognitive science*, *5*(2), 270-282.

Jahnke, H. N. (2008). Theorems that admit exceptions, including a remark on Toulmin. *ZDM*, *40*(3), 363-371.

Kidron, I., & Dreyfus, T. (2014). Proof image. *Educational Studies in Mathematics*, *87*(3), 297-321.

Lakatos, I. (1976). *Proofs and Refutations*. Cambridge: Cambridge University Press.

Larsen, S., & Zandieh, M. (2008). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics*. 67 (3) 205-216.

McGowen, M. A., & Tall, D. O. (2010). Metaphor or Met-Before? The effects of previous experience on practice and theory of learning mathematics. *The Journal of Mathematical Behavior*, *29*(3), 169-179.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in mathematics*, *27*(3), 249-266.

Oehrtman, M. C., Carlson, M. P., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' understandings of function. In M. P. Carlson & C. Rasmussen (Eds.)*, Making the connection: Research and practice in undergraduate mathematics* (pp. 150-171). Washington, DC: Mathematical Association of America.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics*, *66*(1).

Raman, M. (2003). Key ideas: what are they and how can they help us understand how people view proof? *Educational studies in mathematics*, *52*(3), 319-325.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. Schoenfeld, J. Kaput and E. Dubinsky (Eds.)*, Research on Collegiate Mathematics Education, Vol.3,* M.M.A. and A.M.S. 1998, 234-283.

Sandefur, J., Mason, J., Stylianides, G. J., & Watson, A. (2013). Generating and using examples in the proving process. *Educational studies in mathematics*, 83(3), 323-340.

Selden, A., & Selden, J. (1987, July). Errors and misconceptions in college level theorem proving. In *Proceedings of the second international seminar on misconceptions and* *educational strategies in science and mathematics* (Vol. 3, pp. 457-470). Ithaca, NY: Cornell University.

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*, 4-36.

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

Selden, J., & Selden, A. (2015). A perspective for university students' proof construction. In T. Fukawa-Connelly, N. Infante, K. Keene, & M. Zandieh (Eds.), *Proceedings of the 18th Annual Conference on Research in Mathematics Education* (pp. 22-36). Pittsburgh, PA: SIGMAA on RUME.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, *29*(2), 123-151.

Sinclair, N., Watson, A., Zazkis, R., & Mason, J. (2011). The structuring of personal example spaces. *The Journal of Mathematical Behavior*, 30(4), 291-303.

Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017) "Research on the teaching and learning of proof: Taking stock and moving forward." In: J Cai (ed.) *Compendium for Research in Mathematics Education*. Reston, VA: National Council of Teachers of Mathematics.

Swinyard, C., & Larsen, S. (2012). Coming to understand the formal definition of limit: Insights gained from engaging students in reinvention. *Journal for Research in Mathematics Education*, *43*(4), 465-493.

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

Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, 56(2-3), 209-234.