Students' Proof Comprehension via Set-Based Reasoning in a Transition-to-Proof Course

Olivia Bruner
Arizona State University

Kyeong Hah Roh
Arizona State University

Norman Contreras
Arizona State University

Kathleen Melhuish
Texas State University

Lino Guajardo
Texas State University

This study explores Transition-to-Proof (TTP) students' proof comprehension by analyzing written responses to proof intention and validation tasks. We examined students' use of set-based reasoning as they completed these tasks. Our analysis of 28 students' written work found that students used set-based tools for logic to support their reasoning. We also found that their explanations in their responses included discussion of the overall structure, individual lines or components, the mathematical veracity of the claim, and a walk through of the proof-text. We concluded that in these tasks students leveraged set-based reasoning for logic to make sense of the context and claim of the proof, supporting them in their overall proof comprehension.

Keywords: Transition to Proof, Set-Based Reasoning, Proof Comprehension

Reading, writing, and validating proofs are fundamental skills for participation in higher mathematics, so it is vital for undergraduate students to acquire these skills to learn and communicate within advanced mathematics (Tall et al., 2012). Because of this, understanding transition to proof (TTP) students' comprehension of proof, and providing interventions that meaningfully increase proof comprehension for students has been a focus within undergraduate mathematics research (Alcock & Weber, 2005; Selden & Selden, 2003; Dawkins & Zazkis, 2021; Hodds et al., 2014). These studies have demonstrated that interventions can improve students' performance on proof comprehension tasks.

Our study explicitly builds on the work seeking to improve students' proof comprehension by strengthening set-based reasoning for logic within conditional proofs (Dawkins et al., 2023; Dawkins & Roh, 2024). These works have provided evidence of components students deem important in proofs via validation tasks, along with presenting a variety of interventions tailored to this population of TTP students.

Building on the work of Dawkins et al. (2024), we explore how set-based reasoning for logic was evidenced in a class of students' written work on proof intention and validation tasks. While this study shares similarities with existing research on proof validation and set-based reasoning for logic within proof, we focus on analyzing the written artifacts of a TTP class rather than individual interviews or group discussions. In this study, we sought to answer the question: In what ways is set-based reasoning present in transition to proof students' proving activities?

Theoretical Framework and Related Literature

As we focused on student attention in their justifications surrounding proof validity and intention, we adapted existing research in the field and identified three key constructs for this study: *Overall Structure*, *Line-by-Line*, and *Mathematical Veracity*. *Overall Structure* refers to the structure or form of a proof (Selden & Selden, 1995, 2003; Tucci et al., 2023); Mejia-Ramos et al. (2012) emphasize this as a component of holistic proof comprehension. We take *Line-by-Line* to mean student attention related to local aspects of the proof text within their justifications surrounding proof activities, including meanings of terms and individual lines within the proof.

This includes what the existing literature has described as surface features (Selden & Selden, 2003), warrants between lines of a proof-text (Alcock & Weber, 2005), and comprehension of local aspects of a proof (Mejia-Ramos et al., 2012). By *Mathematical Veracity*, we refer to the degree to which students consider a mathematical statement to be proven as true. This concept is characterized by Tucci et al. (2023) as the alignment of the theorem, the proof, and the students' own understanding.

Proof Validation and Intention

Proof comprehension is complex (Mejia-Ramos et al., 2012), so we focus on two task types that provide connected but distinct insights into proof comprehension: Proof Validation and Proof Intention. Proof Validation, determining if a proof text is valid or invalid (e.g., Alcock & Weber, 2005), is a core component of proof comprehension. However, students may not necessarily consider what the proof intends to establish, or determine the Proof Intention (e.g., Roh & Lee, 2024; Tucci et al., 2023). As such, a validation task might overlook the importance of intention unless it specifically asks the student to articulate what the proof sets out to prove. Understanding proof intention involves recognizing the overall structure of the proof and how each part (line-by-line) contributes to what the proof is trying to establish (mathematical veracity). Because of this interrelated nature, both proof validation and proof intention are valuable for gaining insights into how students comprehend proofs.

Set-Based Reasoning for Logic

Finally, this work builds on prior work about set-based reasoning. Set-based reasoning is an approach to logic leveraging sets and set relations as a core tool (Hub & Dawkins, 2018) that has been applied to logic instruction in the TTP environment. To provide context for students' demonstrated proficiency in set-based reasoning, we adopted the codebook and approach of Dawkins et al. (2024), considering normativity and internal consistency of student work in tasks using multiple logical registers (diagrams, set relations, conditionals, etc.). We also developed a new coding set related to student annotations present in the margins, including their logic or set-based understandings of a proof task, detailed in the methodology section.

Methodology

Data Collection

This study is part of a more extensive project (NSF DUE #2141925) aiming at curriculum development for transition-to-proof courses at the undergraduate level. The larger project intends to implement a series of interventions, including set-based reasoning for logic, to support novice provers in their proof comprehension. This paper specifically analyzed part of the data collected from the project's second phase, which implemented the RAMP curricular materials in the classroom at a large Southwestern university.

In order to focus on students' use of set-based reasoning in proof comprehension, we chose tasks in homework and exams that contained problems centered on set-based reasoning, proof validation, and proof intention. We considered the written work of 28 students, and this analysis focused on the midterm, which was the first formal assessment that included proof tasks in the course. The midterm included three multipart tasks that asked for both proof intention and proof validation for provided proof texts with varied contexts and varied normative sufficiency (See Figure 1 below). This midterm occurred in week 7, after one unit of set-based reasoning for logic (weeks 1-3) and one unit of introduction to conditional proofs (weeks 4-6). To provide context for the set-based reasoning of students, we also analyzed two preceding homework assignments from week 4 and week 6.

Task 2

Proof. Let x be any integer that is a multiple of 6.
Then by Definition of Multiples of 6, there exists some integer k_6 such that $x = k_6 * 6$.
Thus, $x = (k_6 * 2) * 3$
Name $k_6 * 2$ as k_3 so as to write $x = k_3 * 3$.
Then k_3 is an integer because it is the product of two integers.
Thus, by Definition of Multiples of 3, x is a multiple of 3. □

Task 3

Proof. Let $\triangle ABC$ be a triangle of a plane with $AB = AC$.
Then there is a point D such that the angle bisector of the angle $\angle A$ intersect the base BC at D.
Then $\triangle ABD$ is congruent to $\triangle ACD$.
Thus, $m\angle B = m\angle C$. □

Task 4

Proof. Let x be an arbitrary integer that is a multiple of 4 and a multiple of 6.
Since x is a multiple of 6, by Definition of "Multiples of 6", there is some integer k_6 such that $x = k_6 * 6$.
Since x is a multiple of 4, either k_6 or 6 must be a multiple of 4.
But 6 is not a multiple of 4. Thus k_6 is a multiple of 4.
Then by Definition of "Multiples of 4", there is some integer k_4 such that $k_6 = k_4 * 4$.
This means

$$\begin{aligned} x &= k_6 * 6 \\ &= (k_4 * 4) * 6 \\ &= k_4 * 24. \end{aligned}$$

Thus, by Definition of "Multiples of 24", x is a multiple of 24. □

Figure 1: Proof Texts Responded to in Midterm Tasks

Data Analysis

We began by developing an initial codebook for attention in proof validation explanations based on the existing literature. The next phase of the data analysis was to refine and finalize the codebooks by adding new codes for the data that did not fit into the initial codebook and grouping the codes into four categories. Through this process, we developed a new code for proof validation emergent from our data*,* *walk-through*, which will be detailed in the results section. In the final phase, we finalized the codebook (see Table 1).

Table 1: Finalized codebook for student explanation in verification/intention tasks

Code	Definition
Line-by-line	Discussion of strengths or weaknesses in the line-by-line argument (e.g., use of definitions, logical or mathematical steps, or arguments or reasoning connecting individual lines of the proof) or show overall concern for the process of the proof argument
Overall Structure	Discussion of the structure or framework of the proof, including explicit attention to the first and last lines of the proof as significant and explicitly stating proof structure (i.e., direct proof, proof by contradiction)
Mathematical Veracity	Discussion of the mathematical correctness of the statement being proven or provides a counterexample demonstrating mathematical falsehood
Walk-through	Evaluation includes a retelling of the proof logic, without other comment or evaluation, a walk-through of the steps taken by the proof-text author

In our initial pass of the data, we also noticed that many students had used the margins of the midterm exam to make notes about the proof validation tasks. We developed a codebook for these notes using open and axial coding (Strauss & Corbin, 1998), as seen in Table 2.

Results

Through this study, we characterized student comprehension of proof-text tasks by considering the written explanations they provided in their responses. We also saw students engage with set-based reasoning in support of their proof comprehension in these tasks. We reflect on the apparent effectiveness of the approaches and variation in methods between the three different presented tasks.

Table 2: Margin Evidence Codebook

Code	Definition
Logical Statements	Annotations that are related to logical reasoning (i.e., logic statement types, generating a conditional statement, or identifying logical equivalence of multiple statements)
Set Based Reasoning	Annotations related to sets and set-based reasoning (i.e., definitions of sets, Euler diagrams, relations between sets)
Examples	Concrete examples or cases considered, such as diagram drawings or arithmetic with example values
Truth Value	Explicit annotation of provided potential conditional as mathematically true or false

Attention in Proof Comprehension

Through iterative refinement of our codebook, we identified four components of the proof text that students attended to when explaining their reasoning in these tasks. These categories are based on student written responses to the prompt "Evaluate if the proof-text properly proves what it attempts to prove. Explain why or why not." These categories are not mutually exclusive; students' written responses often contain multiple parts or arguments, so they often attend to multiple components in the same response.

Overall structure. Many responses contained explicit references to the overall structure of the proof-text. For example, some responses referenced common types of proofs (e.g., proof by contrapositive), content recently introduced in the course. For example, Bailey wrote,

> The proof-text properly proves that for any integer x, if x is a multiple of 6, then x is a multiple of 3. This is a direct proof. (iii) is the contrapositive of the direct proof and would result in the same truth table. However, the proof text does not use the contrapositive format.

Others showed awareness of the overall structure by commenting on the first and last lines of the proof, a technique introduced in a prior curricular activity. In Task 2, Adam wrote:

> The proof text is attempting to prove that for all integers x, if x is a multiple of 6, then x is a multiple of 3. The proof text does indeed prove this properly, as it starts by establishing that x is any integer that is a multiple of 6, breaks it down with the prime factorization, and ultimately concludes with x being shown to be a multiple of 3.

These were not isolated instances. In Task 2, 15/26 responded to the overall structure in their response. This was not consistent across all three tasks, however. 8/26 cited the overall structure in their written explanation for Task 3, and 10/25 cited it in Task 4.

Line-by-line. Students also provided arguments based on surface features or individual components of the proof, which we coded as line-by-line arguments. For example, Tucker's response to Task 3 reads,

> The proof-text does not properly prove what it attempts to because there is a gap in justification between line 2 and line 3. The proof-text does not point out the consequence of the angle bisector and quickly jumps to line (3).

Arguments in this category were cited as justification for the proof's validity. For example, Evan's response to Task 2 cited the internal components of the proof as the reason the proof proves what it attempts to, saying: "Yes, because it logically states and justifies each step using either algebra or a definition that ultimately arrives at the desired conclusion."

Overall, 13/26 responses to Task 2, 19/26 to Task 3, and 10/25 to Task 4 included a line-by-line argument. Some responses included both line-by-line and overall structure components; 14/77 responses were coded as both. This shows that students attending to overall structure do not necessarily cease to use line-by-line reasoning in their validation explanations.

Mathematical veracity. Some explanations demonstrated attention to the correctness of the statement being proven, which we called mathematical veracity of their answer a claim. For example, John response reads:

> The proof tries to prove that multiples of 4 and 6 are always multiples of 24 but doesn't do it properly because they neglected the case where multiples of 4 and 6 are also every other multiple of 12 as well.

In Tasks 2 and 3, with 2/26 and 1/26 responses coded as attending to mathematical veracity, and in Task 4 8/25 responses cited mathematical veracity. This is not surprising; compared to the previous two categories, mathematical veracity doesn’t directly connect to a recent curricular activity. Fewer responses coded as citing mathematical veracity likely do not have to do with what students think is important but rather what recent course activities have primed them to include in responses to proof-text tasks.

Walk-through. While most student explanations analyzed contained arguments we expected to see based on previous literature, we also noticed a new focus. Some responses included a retelling of the proof itself, or a walk-through of the reasoning of the proof as justification for the proof properly proving what it set out to prove. For example, Ben's response to Task 3 contains a walk-through of the proof-text:

> The proof-text attempt to prove that for any triangle ABC, if $AB = AC$, then $m\angle B = m\angle C$. The prove does not fully prove the prove properly. The statement "Then triangle ABD is congruent to triangle ACD" is not properly justified. The proof start state how the point D from the angle bisector creates line AD. Then, since $AB = AC$, $\angle CAD$, and $\angle DAB$ are equivalent by definition of angle bisector, and $AD = AD$ triangle ABD is congruent to triangle ACD. The proof is mostly correct and just would benefit from some clarity.

In most cases the walk-through was included along with another category of explanation. Walk-through was the only code applied to a written response only 4/19 of the times it was used.

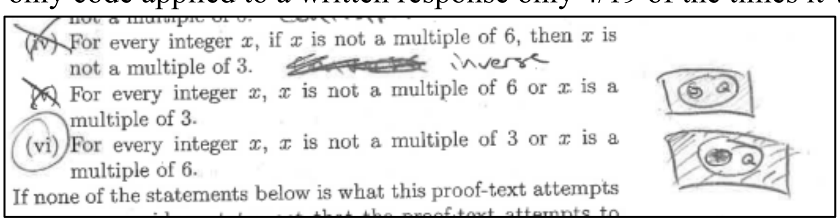

Figure 2: Example of student set-based margin annotations

Engagement in Set-Based Reasoning

In the process of our analysis, we observed margin annotations present (see Figure 2) on all three tasks, including a variety of annotations, some of which were set-based reasoning tools that had been presented earlier in the course. We had observed general proficiency, though not without error, with set-based reasoning for logic in the work of this same class on prior homework assignments, using the same approach and codebook as Dawkins et al. (2024), but due to space constraints, we focus only on the set-based tools students used on these midterm tasks. However, it was because of this context that we attended to these margin annotations.

Evidence of set-based reasoning in proof comprehension contexts. Since the midterm questions did not require students to show their work in the question stem, we can conclude that these annotations were added because they were of value to the student in their problem-solving process on the task. Based on these observations, we developed the codebook in Table 3. These margin annotations demonstrate that students found set-based tools for logic useful as they solved proof intention and verification tasks (see Table 2).

Table 3. Number rates of occurrence in margins of midterm tasks

Task	Set-based	Logic	Example	Truth
2	12/26	14/26	1/26	1/26
3	4/26	6/26	21/26	2/26
4	3/25	12/25	4/25	1/25

The margin work also indicated attention to the logical reasoning skills previously introduced as part of the RAMP curriculum, showing student uptake of those concepts and skills. This is unsurprising, as these concepts were freshly introduced and directly associated with proofs in the way they were presented. However, this is also some indication of adoption, as students found these concepts valuable enough to make personal notes to support their proof comprehension on these tasks.

Diversity of use amongst tasks. The margin annotations were not uniformly present across the three tasks, as evidenced in Table 3. Two components likely contributed to this. First, the mathematical context of the task seemed to be associated with the margin notes used; specifically, the use of examples was much higher on Task 3, the only geometry-based task. It seems likely to us that students could run examples of the divisibility context in their mind's eye but that the more spatially challenging context of geometry led to the higher use of examples as a written aid. Also, set-based tools appeared at a much higher frequency on Task 2 than on Tasks 3 and 4. However, all students who used set-based tools on Task 3 or Task 4 had also used them on Task 2. This suggests that because of the similar setups of these three problems, some students used set-based tools to help them conceptualize the first task but then transferred their conclusions to the next two questions that were of a similar form, without a need to reconstruct the annotations they had previously made. Also, the question directly preceding Task 2 on the assessment asked about equivalent sets; it is possible that this prior task primed set-based thinking in the students. However, we posit that the students would not have used set-based annotations if they found them fully irrelevant to the task at hand.

Our analysis of the students' work on these tasks, in light of our observations of their work on prior set-based homework, suggests that while set-based reasoning might not be trivial to the students, it was accessible to them. In particular, the students' use of set diagrams, notations, and terminologies in their margin scratch work indicates that students found this set-based reasoning approach beneficial for reading and comprehension of proof texts on these types of tasks.

Normative Proof Validation

While we avoid making statistical claims past describing the proportion of the responses that fit in each category, as this was not a representative sample, we did see patterns in response categories of attention, margin note types, and the normativity of the response to the question of the proof text's properness, which we considered a proof validation task. These observations, we believe, provide insights as to how these techniques connect to proof comprehension broadly.

Overall structure and mathematical veracity in validation explanations. Student attention to overall structure in their explanation of reasoning did not consistently co-occur with a normative evaluation of the proof text in these three tasks. In Task 2, all 15 responses citing overall structure correctly validated the proof as proper, and 8/9 overall structure responses validated the proof as proper in Task 3. However, Task 4 tells a different story. 9/10 responses coded for overall structure declared the proof "proper" even though its claim is mathematically false. This suggests that the responses that attended to overall structure, while relatively successful at the proof intention component of the task, were not consistently normatively validating the conditional they successfully identified.

In Task 4, many responses that correctly identified the proof as improper were coded as mathematical veracity, (see Table 1 for definition). For example, John responded:

> The proof tries to prove that multiples of 4 and 6 are always multiples of 24 but doesn't do it properly because they neglected the case where multiples of 4 and 6 are also every other multiple of 12 as well.

In Task 4, 8/25 responses cited mathematical veracity, among which 7/8 correctly declared the proof improper in proving what it attempts to prove. While attention to overall structure in the response seemed to help students complete proof intention tasks normatively, mathematical veracity most frequently co-occurred with a normative declaration of a proof as invalid.

Conclusion/Discussion

Through our analysis of these two interrelated proof comprehension task types, proof intention and proof validation, we found evidence of students' use of set-based reasoning tools from their TTP course. Students used overall structure arguments in their explanations of proof validation and included components of set-based reasoning and logical structures in their margin work on the tasks. Because margin work was not a required or graded component of the assessment, we take this to mean that the students found these tools relevant to the presented proof comprehension tasks. It is worth noting that these set-based tools were not self-emergent; the intention of the course was to teach intro to proofs using set-based reasoning for logic.

Many of the aspects of proof we saw students attend to were consistent with those described in prior work, including student focus on proof structures (Selden & Selden, 2003; Tucci et al., 2023), surface feature or line-by-line reading (Selden & Selden, 2003; Dawkins & Zazkis, 2021), and a focus on mathematical veracity of the claim made (Tucci et al., 2023). This demonstrates that these categories continue to be evidenced in full class as demonstrated on written proof comprehension tasks. However, we did observe a new component included in the written explanations, that of a walk-through of the proof text. Though this was a small set of tasks, it does seem from these written responses that walk-through occurs as a way for students to say it's right because it's right, since among these artifacts, a claim of improperness tended to include identifying a failure within the proof, leading to a code of line-by-line instead. It is also possible that this retelling action is a byproduct of the novelty of the context and the task; students may be retelling themselves the argument of the proof as they work to explain their reasoning. However, this is a conjecture, and so we believe that more investigation into this novel code of walk-through as an explanation is warranted.

Acknowledgments

This research was funded by NSF DUE #2141925. All opinions are those of the authors and do

not necessarily represent the views of the National Science Foundation.

References

Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *The Journal of Mathematical Behavior*, *24*(2), 125–134. https://doi.org/10.1016/j.jmathb.2005.03.003

Dawkins, P. C., & Roh, K. H. (2024). Unitizing Predicates and Reasoning About the Logic of Proofs. *Journal for Research in Mathematics Education*, *55*(2), 76–95. https://doi.org/10.5951/jresematheduc-2020-0155

Dawkins, P. C., Roh, K. H., Bruner, O., & Gonzalez, M. (2024). Student use of three representations of set relationships to reason about logic in undergraduate transition to proof courses. *Proceedings of the 45th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Cleveland, OH.*

Dawkins, P. C., Roh, K. H., & Eckman, D. (2023). Theo's reinvention of the logic of conditional statements' proofs rooted in set-based reasoning. *The Journal of Mathematical Behavior*, *70*, 101043. https://doi.org/10.1016/j.jmathb.2023.101043

Dawkins, P. C., & Zazkis, D. (2021). Using moment-by-moment reading protocols to understand students' processes of reading mathematical proof. *Journal for Research in Mathematics Education*, *52*(5), 510–538. https://doi.org/10.5951/jresematheduc-2020-0151

Hodds, M., Alcock, L., & Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education*, *45*(1), 62–101. https://doi.org/10.5951/jresematheduc.45.1.0062

Hub, A., & Dawkins, P. C. (2018). On the construction of set-based meanings for the truth of mathematical conditionals. *The Journal of Mathematical Behavior*, *50*, 90–102. https://doi.org/10.1016/j.jmathb.2018.02.001

Mejia-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, *79*(1), 3–18. https://doi.org/10.1007/s10649-011-9349-7

Roh, K. H., & Lee, Y. H. (2024). Undergraduate students' logical consistency in mathematical thinking: Implications for teaching and learning. *PRIMUS*, 1–13. https://doi.org/10.1080/10511970.2024.2379370

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, *34*(1), 4–36. https://doi-org.ezproxy1.lib.asu.edu/10.2307/30034698

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. Educational Studies in Mathematics, 29(2), 123–151. https://doi.org/10.1007/BF01274210

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory, 2nd ed* (pp. xiii, 312). Sage Publications, Inc.

Tall, D., Yevdokimov, O., Koichu, B., Whiteley, W., Kondratieva, M., & Cheng, Y. H. (2012). Cognitive development of proof. *Proof and proving in mathematics education: The 19th ICMI study*, 13-49.

Tucci, A., Dawkins, P., Roh, K., Eckman, D., & Ruiz, S. (2023). *Student explanations and justifications regarding converse independence*. Proceedings of the Annual Conference on Research in Undergraduate Mathematics Education.