Analyzing an Instructor's Oral Follow-Up Assessment Strategies in an Introductory Proof Course

Nurul Wahyuni
Davis & Elkins College

David Miller
West Virginia University

This study investigates the characteristics of an instructor's communication in one-on-one oral assessments for an introduction-to-proof course and how they impact student learning. It examines the instructor's strategies for evaluating the depth of students' knowledge. Using thematic analysis, the study examines patterns in question types and their relation to revision quality and presentation fluency. Findings reveal that well-revised proofs prompted questions about comprehension and logical structure. For revisions needing improvement, questions guided students to establish correct proofs before addressing comprehension. The study also explores teaching elements resulting from the instructor's assessment practice using commognitive theory. These findings contribute to assessment practices, particularly in revision and oral assessment for introduction-to-proof courses. By highlighting communication's role in student learning, this research advances alternative assessment understanding in higher education mathematics.

Keywords: revision, oral assessment, objective-based grading, commognition, proof feedback

One crucial element of teaching practice revolves around the creation and execution of assessment methods. Gaining insight into how an instructor carries out his/her teaching practices can offer valuable insights into the implementation of a course. Given the significant role mathematical proofs play in the field of mathematics, evaluating proofs forms an integral aspect of teaching practice. Prior research has delved into how mathematics instructors assess students' written proof submissions using conventional point-based systems (Miller et al., 2018; Moore, 2016). This study aims to extend our understanding by exploring oral assessment methods, specifically focusing on objective-based grading (OBG) in an introduction-to-proof course.

The course that is being studied in this paper implemented oral assessments to evaluate proof revisions under the implementation of OBG. In OBG, student performance is typically assessed as pass or fail, based on near-perfect or perfect solutions, with minor errors sometimes overlooked. However, making a judgment about how serious a mistake is can be a challenging task when evaluating mathematical proofs (Miller et al., 2018; Moore, 2016). Oral assessment, however, has the advantage of clarifying whether a small mistake was indeed insignificant to the overall mastery of the material through follow-up questions.

Studies highlight the advantages of oral assessments for evaluating students' proof skills (Soto-Johnson & Fuller, 2012; Stylianides, 2019) and reveal that when students orally present their proofs, they often offer more comprehensive explanations, resulting in stronger proof outcomes (Soto-Johnson & Fuller, 2012; Stylianides, 2019). Other studies have discussed the implementation and outcomes of the use of OBG (e.g., Cooper, 2020; Iannone & Vondrová, 2023; Prasad, 2020; Zimmerman, 2020), and have discussed the benefits of oral assessment (e.g., Iannone et al., 2020; Iannone & Simpson, 2012, 2015). However, there has been limited research on oral assessment with OBG and this study will inform readers about the instructor's communicative approach to the implementation of oral assessment in an introduction-to-proof class using OBG. The main research goals of this study are: (1) What are the discursive characteristics of the instructor's oral communication during oral follow-ups when grading students' proof productions? and (2) What is the significance of the instructor's communication practices in the oral follow-up assessment?

Literature Review

Objective-based Grading

The traditional grading system still dominates higher education mathematics in the U.S., employing partial credit for each problem submitted (Nilson, 2015). There is, however, an alternative grading approach known as OBG, also referred to as mastery-based grading or standard-based grading. Under this system, the course is organized around a set of specific objectives that students must successfully achieve. Students are provided with multiple chances to demonstrate their comprehension for each objective and any initial failure to pass an objective does not impact their overall grade; only the final outcome is counted (Campbell et al., 2020).

Studies consistently demonstrate that OBG offers numerous advantages for both students and instructors (Knight & Cooper, 2019; Linhart, 2020; Zimmerman, 2020). Students benefit from reduced testing anxiety through multiple attempts, learning from mistakes, fostering autonomy, and prioritizing learning over points. This approach can significantly impact the depth of their understanding and intrinsic motivation (Linhart, 2020). Instructors, on the other hand, find satisfaction in streamlined grading processes without the need for complex rubrics, appreciate their students' increased effort and improved quality of work, gain insight into students' capabilities and understanding through objective-based assessment, and experience enhanced purposefulness in their teaching (Knight & Cooper, 2019; Prasad, 2020).

Proof Evaluation Practices

Teaching goes beyond classroom instruction; as Moore (2016) noted that when professors assess students' proof productions, they effectively engage in teaching by making corrections, providing comments, and discussing students' thought processes to enhance their proof-writing skills. Evaluating proof productions is a challenging aspect of teaching, influenced by individual judgments of the seriousness of proof errors, student comprehension, and perceived abilities, leading to considerable scoring variability among professors, especially in traditional point-based grading (Miller et al., 2018). However, OBG minimizes this variation, resulting in more consistent grading outcomes, except in cases when the seriousness of proof errors is disputed.

Mathematicians grade students' proofs based on several aspects, such as logic, clarity, fluency, understanding (Moore, 2016), and perceived ability (Miller et al., 2018). Some professors also focus on linguistic and academic rules, which students often overlook (Lew & Mejía-Ramos, 2019). Surprisingly, students and professors have contrasting views about language conventions on proofs. For instance, students, regardless of language skills, do not realize the importance of correct academic English in proofs. Mathematicians emphasize complete sentences, with most deducting points for incomplete ones, while only a few students find this unconventional. In addition, professors and students have varying opinions about overusing variables in proofs (Lew & Mejía-Ramos, 2019).

Oral assessments may have distinct evaluation criteria compared to written assessments, especially when students prepare written solutions in advance without supervision. In oral assessments, follow-up questions are crucial for evaluating a student's understanding, making the grading criteria more complex (Iannone & Simpson, 2012). This study aims to contribute to mathematics education literature by shedding light on the evaluation of students' proof productions in oral assessments, addressing a gap in the existing literature. It will provide insights for both researchers and professors in higher education mathematics on the practice of oral assessments and proof feedback in higher education mathematics.

Theoretical Framework

From a commognitive perspective, human actions fall within specific discourses, each defined by its objects, communication methods, and participant rules, thereby establishing distinct communities of communicators (Sfard, 2008). For example, discussions centered on mathematics form a mathematical discourse. Within mathematics, a discourse is built upon structured activities known as mathematical routines, which Sfard (2008) defines as "a set of metarules that describe a repetitive action" (p. 208). These routines serve the purpose of generating narratives about mathematical concepts and objects.

Kontorovich (2021) introduced the concept of the didactical discourse on proof (DDP) to encompass the discourse of the teaching and learning of mathematical proofs. DDP has two key aspects: the pedagogical aspect and the mathematical aspect (Kontorovich, 2021). For the pedagogical role, DDP guides educators on what and how to teach in a proof course. For the mathematical aspect, participants must adhere to community-accepted rules for mathematics when discussing mathematical proofs. The primary goal of this research is to uncover patterns of DDP within the context of assessing mathematical proof productions through oral assessments.

He also examined routines in feedback practices using DDP and the commognitive framework to analyze proof feedback in a graduate topology course, focusing on comments and point deductions on graded assignments. He found that fully-scored proofs received comments related to mathematical concepts (mathematizing), while reduced-score proofs often received comments addressing the student personally (subjectifying). Point deductions were always linked to issues concerning the proof's idea, but not its representation (Kontorovich, 2021).

Viirman (2014, 2015) explored routines among mathematical instructors during lectures. Viirman (2014) discovered variations in lecture routines, including construction routines (common to all instructors, involving concept definitions and diverse examples) and substantiation routines (related to theorem proving, with varying emphasis). Viirman (2015) listed three key didactical routines: explanation routines for clarifying concepts, motivation routines to engage students, and question posing routines (featuring rhetorical questions).

The didactic practices of mathematics instructors remain relatively unexamined. There is very little (or no) existing research that delves into the routines involved in the execution of oral assessments within higher education mathematics courses. This study aims to contribute to the body of literature in mathematics education by exploring the routines employed by mathematics instructors when conducting oral assessments to evaluate mathematical proofs.

Methods

Context

The research was conducted in an introduction to proof course, involving four students and an instructor with pseudonyms of Amy, Clara, Dana, Gina, and Dr. Jones. The study focused on evaluating proof productions on quizzes, following a specific process. Initially, students completed a closed-book, in-person written quiz in which each learning objective was graded as a success, growing, or not assessable. Only a grade of success counted as passing and contributed to the final grade. For students who did not obtain success on certain learning objectives on the quiz, they had the opportunity to revise their solutions and present them in a one-on-one virtual oral follow-up assessment within a week of receiving feedback. During this oral assessment, Dr. Jones could ask additional questions related to the presented proofs.

Data Corpus

The data for this study was comprised of: (a) original proof productions by students, (b) feedback from the instructor, Dr. Jones, (b) students' revised written proof productions, (d) recordings of the oral follow-up meetings, and (e) interviews with Dr. Jones. The study concentrated on proof constructions, assessed on Quizzes 2 through 5. Four participants (Amy, Clara, Gina, and Dana) presented 4, 10, 10, and 3 revised proofs, respectively, in some of the oral follow-up assessment meetings.

The interviews with Dr. Jones were conducted twice through Zoom using a semi-structured format, covering various themes. The main topics discussed during the interviews included Dr. Jones' teaching background, course structure, quiz assignment procedures, oral follow-ups, written feedback practices (both general and specific), and the implementation of mastery grading. The first interview with Dr. Jones took place at the end of the semester, and the second interview occurred approximately three months after the course had concluded, focusing more on the implementation and outcomes of mastery grading in the course.

Analysis

The types of follow-up questions were analyzed in relation to the quality of revised proof productions. The revised productions were categorized into one of the following: successful revised proof production with no improvement or correction suggested by Dr. Jones - *coded as P1*; revised proof production that is correct but could be improved by adding more information, such as additional details or explanations, or by refining some phrases - *coded as P2*; revised proof production that has one or more mathematical errors – *coded as P3*.

The oral follow-up recordings were the main data for this study. The thematical analysis coding was conducted using MAXQDA software for qualitative analysis. Because of our interest in depicting Dr. Jones' communication routines, the codes were developed to emphasize the aspects of communication in Dr. Jones' phrases. We particularly focused on capturing on the type of communication, rather than the specific content of the communication itself. For instance, when Dr. Jones inquired, "Why did you change the quantifier?" the code used is "question on changes" rather something related to quantifier. This approach allows us to center the analysis on the communication aspects rather than the mathematical aspects.

Results

The analysis revealed that Dr. Jones employed two distinct approaches when interacting with students based on the quality of the revised proof productions (P1, P2, or P3): the routines when encountered with correct narratives and reconstruction of an alternate narrative.

Routines when Encountered with Correct Narratives

There are two possible cases of Dr. Jones' actions when encountered with correct proof productions: the proof was instantly endorsed without any follow-up questions or there was an examination of the students' comprehension of the proof presented. When questions were posed, they were usually centered around logical aspects of the proof or the rationale for changes in the proof.

From a total of 27 proof presentations in the oral follow-up sessions, 11 of them were categorized as P1, which means the revised proofs had no issues. This categorization was based on the nonexistence of Dr. Jones' comments for disapproval or comments on improvement. Among these 11, there were five performances that received successes instantly. During the interview with her, Dr. Jones pointed out two important aspects describing the circumstances that do not require follow-up questions. The first aspect is the authenticity of writing. This is important because the nature of the assessment allows the students to access any resources available to them. However, they are expected to be able to write it in their own words. The second aspect is the fluency of their verbal communication in which the presentation of the proof should sound very natural. Therefore, a good quality of mathematical verbal communication is valued not only in terms of its validity, but also the fluency of the explanation part.

In most cases, error-free revised proof productions were subject to scrutiny by Dr. Jones. This was done to ascertain whether students had indeed learned and understood the material that they presented. Dr. Jones mentioned that “the oral follow-up gave me that option to figure out if they just copied it from somewhere or if they had really learned something.” The routines of Dr. Jones implemented when she encountered an error-free proof productions can be distinguished into two types of inquiry: Question/prompts related to the logical structure of the proofs and questions/prompts related to the changes from the originals to the revised versions of the proofs.

Question/prompts related to the logical structure of the proofs. Dr. Jones collected information about the students' logical structures through implicit and explicit methods. In the explicit approach, Dr. Jones would request students to explain the reasoning behind their proof structures. Here's an illustration of how this explicit examination of the proof's logic works.

Gina revised her proof for the following objective: “I can construct a correct proof of a conditional statement involving sets.” The statement to prove was as follow: Let A, B, V, and W be sets. Prove that if $V \subseteq A$ and $W \subseteq B$, then $V \cap W \subseteq A \cap B$. After Gina presented her revised proof for this objective, Dr. Jones asked her the following question, “So, what’s the key word that tells us ultimately that definition of subset? That key word that’s allowing us to conclude V intersect W is a subset of A intersect B?” Gina replied, “The key word?”, to let Dr. Jones know she did not understand the question. Dr. Jones then clarified by stating, “So, somewhere in your proof you’re saying by definition of subset. So, what is that definition of subset? Where’s that in your proof that it’s allowing you to make that conclusion?”

Dr. Jones questioned Gina’s understanding of how the definition of "subset" influenced her proof's logic. Gina struggled to identify this crucial element in her proof. After some discussion, Dr. Jones clarified further, “that's the thing [circling the word *implies*] that actually allows you to make this conclusion [underlying $V \cap W \subseteq A \cap B$] at the end. Because if you had written that as an and then you will be talking about intersection […]. So, this is [pointing out the word *implies*], this word is what's telling us that you're using this definition [the definition of subset].”

In the implicit approach, Dr. Jones questioned the student about a statement's definition without directly connecting it to the proof. This allowed Dr. Jones to assess the consistency between the applied definition and the student's narrative. In the following discussion, Dana presented her revised proof for the statement: “If $A \subseteq B$, then $A - C \subseteq B - C$". After Dana presented the proof, Dr. Jones asked, “Alright. Here's two statements [wrote "$x \in A$ and $x \in B$" and “if $x \in A$, then $x \in B$"]. Are those two statements the same or are they different?” Although Dana accurately revised a proof that relies on the concept of a subset, she struggled to correctly identify the definition of a subset. Dr. Jones then provided the correct information in which she explained, “So, this one is the definition of subset [pointing to the second one]. We're talking about we've got a set A and it's inside of some larger set B. But the first one, we're talking about our two sets don't necessarily have to be nested and things are just in between them. So, this is the definition of intersection [pointing out to the first one].” This exchanged suggests that Dana might not have had a complete grasp of the logical structure underlying her written proof.

Dr. Jones did not immediately allow Dana to falter; instead, she offered her another opportunity through a series of different questions. Dr. Jones then tried to delve into Dana's logical structure more directly on another proof by asking, "How do you know to start with x is an element of A minus C?" During the second round of questioning, Dr. Jones found Dana's response satisfactory and deemed it a success and choose to provide further clarification after acknowledging the student's success, emphasizing the answers to the questions.

Questions/Prompts Related to Changes Made. On many occasions, success in the follow-up meeting requires students to not only construct a correct proof but also explain why the initial proof was rejected. This ability is crucial for meeting the learning target. Failing to articulate the errors in the original proof, even if the revised one is correct, can still result in failure.

This example illustrates how Dr. Jones probed a student's understanding on the changes she made from the original to the revised proof. Gina presented her revised proof for the following problem: "Let A, B, and C be sets. Prove that $A \times (B \cap C) = (A \times B) \cap (A \times C)$." In Gina's initial proof, she had notational problems, like omitting parentheses when denoting an element in the Cartesian product. However, she successfully rectified this issue in her revised proof, prompting Dr. Jones to pose the question, "Also, things like parentheses around that x, y, why is that important?" A student who presented a revised proof correctly may still not achieve success if she cannot provide a clear explanation for why the alterations were required.

Reconstruction of an Alternative Narrative. When a student presents a correct proof production, Dr. Jones may make a point to highlight an alternative version. The alternative version can involve a minor modification or substantial changes to student's version of the proof. Among the data collected, there were three instances when Dr. Jones constructed alternative proof productions. The construction process was conducted through a dialogue with the student in which Dr. Jones guided the process. For example, Dr. Jones illustrated that a proof by cases (with even and odd integers cases) could be simplified to a straight-forward direct proof.

Routines when Encountered with a Flawed Narrative

When a student presented a revised proof that had an error, Dr. Jones provided her with prompts and questions to rectify the issue. The observed flaws in students' revised proofs fell into two categories: those needing additional information or details (P2) and those that were mathematically or logically incorrect (P3).

Eliciting Missing Information. When gaps in the proof arose from missing information, Dr. Jones prompted students to fill in that missing information. In one of Clara's revised proofs, there was a lack of crucial details regarding the requirement for a valid fraction, specifically that the denominator should not be zero. The excerpt below shows how Dr. Jones guided Clara to provide more details about the object's properties used in her proof. Dr. Jones began with the question, "So, those integers a and b that you're using to make up as a rational [number], is there any other condition that should be there?" Clara replied, "Yeah, that they are co-prime, right?" Dr. Jones then stated, "Co-prime, I'm less concerned about that" and Clara exclaimed, "Okay, well, b is not zero." Dr. Jones replied encouraging, "Yeah, that one is kind of important."

Exploration of Errors. When a student presented a revised proof that had some mathematical errors, Dr. Jones would invite the student to discuss the revision. An example of this was shown in Amy's revised proof presentation of the following statement: "Prove: For all integers x and y, if $x + y$ is odd, then $x \neq y$." In her revised proof, Amy initially wrote the contrapositive statement as, "There exist some integers x and y such that $x = y$." In response to this alteration, Dr. Jones posed the question, "Why did you change the quantifier?" This inquiry prompted Amy to reconsider her choice regarding the quantifier. After careful thought about the quantifier and statement, Amy concluded that altering the quantifier was not necessary when applying the contrapositive. Dr. Jones confirmed the validity of Amy's explanation regarding the equivalence of the contrapositive statement. In summary, when a student struggled fixing a proof error, Dr. Jones would start the proof reconstruction process, which might include verbal guidance or an explicit conversation-based reconstruction.

Summary and Discussion

Dr. Jones' approach to assessing oral proof presentations combined flexibility and thoroughness. Dr. Jones occasionally granted success without follow-up questions for authentic and fluently presented valid proofs. However, she frequently asked follow-up questions to confirm understanding, even when proofs appeared correct These queries focused on the logical structure or the reasoning behind revisions. These inquiries helped gauge student comprehension, considering any resources were allowed. When alternative solutions are available, Dr. Jones would highlight this and occasionally provide explanations or walk students through the proof's reproduction. She also gave students opportunities to correct errors through questions or prompts to extract missing information or initiated discussions to encourage self-identification and correction of the proof errors.

Oral follow-up questions revealed a key insight: constructing a correct proof does not necessarily imply a deep grasp of the foundational concepts it relied on. These thought-provoking questions played a pivotal role in preparing students to actively participate in discussing the "why" behind the proof, delving into the rationale and reasoning that support their revisions and the validity of their proofs. This active engagement, in which students defended their arguments, fostered an increased awareness of the meta-discursive rules inherent in constructing a mathematical proof. Encouraging students to articulate and justify their thought processes helped them understand the fundamental principles and logical connections in mathematical proofs, improving their overall communication skills in this context.

Dr. Jones' assessment method aligns well with the conceptualized commognitive process "of becoming able to have mathematical communication" (Sfard, 2007, p. 573). Through revisions and oral presentations, students not only share their proofs but also actively engage in the written and oral communication to deepen their understanding of proofs. This study highlights the effectiveness of oral follow-up assessments in improving communication skills and shifting the focus from acquiring knowledge to expressing mathematical concepts sufficiently.

The act of verbally explaining ideas during the oral follow-up sessions allows students to become involved in an authentic mathematical conversation which promotes mathematical discourse and think critically when responding to questions related to their proof productions. Moreover, through follow-up questions, the students will learn that they need to be responsible to defend every statement presented in their proof narrative. This promotes the metalevel learning on substantiation in their mathematical proofs.

As a final remark, we observed that this assessment method also offers a valuable teaching and learning platform, allowing Dr. Jones to guide students in error identification and proof reconstruction on a one-to-one basis, which is highly effective for learning. Beyond assessing proof productions, the oral assessment encourages mathematical discourse and discussion, promoting deeper understanding and knowledge enrichment. Furthermore, it prioritizes proof comprehension over construction, with discussions shifting from correcting construction errors to probing comprehension when students present correct proofs, thereby facilitating a more comprehensive assessment of students' mathematical understanding.

References

Campbell, R., Clark, D., & OShaughnessy, J. (2020). Introduction to the Special Issue on Implementing Mastery Grading in the Undergraduate Mathematics Classroom. In: Taylor & Francis.

Cooper, A. A. (2020). Techniques grading: Mastery grading for proofs courses. *PRIMUS*, *30*(8-10), 1071-1086.

Iannone, P., Czichowsky, C., & Ruf, J. (2020). The impact of high stakes oral performance assessment on students' approaches to learning: a case study. *Educational Studies in Mathematics*, *103*(3), 313-337.

Iannone, P., & Simpson, A. (2012). Oral assessment in mathematics: implementation and outcomes. *Teaching Mathematics and Its Applications: International Journal of the IMA*, *31*(4), 179-190.

Iannone, P., & Simpson, A. (2015). Students' views of oral performance assessment in mathematics: straddling the 'assessment of'and 'assessment for'learning divide. *Assessment & Evaluation in Higher Education*, *40*(7), 971-987.

Iannone, P., & Vondrová, N. (2023). The novelty effect on assessment interventions: a qualitative replication study of oral performance assessment in undergraduate mathematics. *International Journal of Science and Mathematics Education*.

Knight, M., & Cooper, R. (2019). Taking on a New Grading System: The Interconnected Effects of Standards-Based Grading on Teaching, Learning, Assessment, and Student Behavior [research-article]. *https://doi.org/10.1177/0192636519826709*. https://doi.org/10.1177_0192636519826709

Kontorovich, I. (2021). Minding mathematicians' discourses in investigations of their feedback on students' proofs: a case study. *Educational Studies in Mathematics*, 1-50.

Linhart, J. M. (2020). Mastery-based testing to promote learning: Experiences with discrete mathematics. *PRIMUS*, *30*(8-10), 1087-1109.

Miller, D., Infante, N., & Weber, K. (2018). How mathematicians assign points to student proofs. *The Journal of Mathematical Behavior*, *49*, 24-34.

Moore, R. C. (2016). Mathematics professors' evaluation of students' proofs: A complex teaching practice. *International Journal of Research in Undergraduate Mathematics Education*, *2*(2), 246-278.

Nilson, L. B. (2015). *Specifications grading: Restoring rigor, motivating students, and saving faculty time*. Stylus Publishing, LLC.

Prasad, P. V. (2020). Using revision and specifications grading to develop students' mathematical habits of mind. *PRIMUS*, *30*(8-10), 908-925.

Sfard, A. (2007). When the rules of discourse change, but nobody tells you: Making sense of mathematics learning from a commognitive standpoint. *The journal of the learning sciences*, *16*(4), 565-613.

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. Cambridge University Press.

Soto-Johnson, H., & Fuller, E. (2012). Assessing proofs via oral interviews. *Investigations in Mathematics Learning*, *4*(3), 1-14.

Stylianides, A. J. (2019). Secondary students' proof constructions in mathematics: The role of written versus oral mode of argument representation. *Review of Education*, *7*(1), 156-182.

Viirman, O. (2014). The functions of function discourse–university mathematics teaching from a commognitive standpoint. *International journal of mathematical education in science and technology*, *45*(4), 512-527.
Viirman, O. (2015). Explanation, motivation and question posing routines in university mathematics teachers' pedagogical discourse: a commognitive analysis. *International Journal of Mathematical Education in Science and Technology*, *46*(8), 1165-1181.
Zimmerman, J. K. (2020). Implementing standards-based grading in large courses across multiple sections. *Primus*, *30*(8-10), 1040-1053.