Int. J. Res. Undergrad. Math. Ed. (2024) 10:284–317
https://doi.org/10.1007/s40753-022-00200-0

Lessons Learned About Incorporating High-Leverage Teaching Practices in the Undergraduate Proof Classroom to Promote Authentic and Equitable Participation

Kathleen Melhuish[1] · Paul C. Dawkins[1] · Kristen Lew[1] · Sharon K. Strickland[1]

Accepted: 30 September 2022 / Published online: 4 November 2022

Abstract

In recent years, professional organizations in the United States have suggested undergraduate mathematics shift away from pure lecture format. Transitioning to a student-centered class is a complex instructional undertaking especially in the proof-based context. In this paper, we share lessons learned from a design-based research project centering instructional elements as objects of design. We focus on how three high leverage teaching practices (HLTP; established in the K-12 literature) can be adapted to the proof context to promote student engagement in authentic proof activity with attention to issues of access and equity of participation. In general, we found that HLTPs translated well to the proof setting, but required increased attention to navigating between formal and informal mathematics, developing precision around mathematical objects, supporting competencies beyond formal proof construction, and structuring group work. We position this paper as complementary to existing research on instructional innovation by focusing not on task trajectories, but on concrete teaching practices that can support successful adaption of student-centered approaches.

Introduction

In the United States, there has been a substantial push for undergraduate mathematics to move away from a traditional lecture model (Abell et al., 2018; Saxe & Braddy, 2015). To support these efforts, there are a number of research-based curricula designed to center student thinking (e.g., Larsen et al., 2013); however, there remain a number of open questions related to instructional implementation strategies and the nuances of the proof setting. Moreover, recent results suggest that inquiry-oriented curricula can inadvertently produce inequitable outcomes (Johnson et al., 2020). Researchers have

✉ Kathleen Melhuish
melhuish@txstate.edu

1 Department of Mathematics, Texas State University, 601 University Drive, San Marcos, TX 78666, USA

conjectured that these inequities may result from a number of sources including inequitable participation where certain students may take on more substantial roles in mathematical activity. Such conjectures align with well-documented status issues that emerge in group work in K-12 mathematical settings (e.g., Esmonde, 2009) and preliminary work in the proof-based setting (e.g., Brown, 2018).

Much of the design-based research work at the advanced undergraduate level relies on content-driven rather than participation-driven design heuristics. By this we mean the overarching objective is for students to reinvent concepts, theorems, and algorithms (e.g., Larsen, 2013; Rasmussen & Kwon, 2007; Wawro et al., 2012). However, disciplinary practices often undergird the reinvention processes (such as in the analysis of Larsen & Zandieh, 2008; Rasmussen et al., 2015). In our design project, we fore fronted participation in disciplinary practices as the primary student activity goal. In particular, we aimed to engage students in disciplinary practices that support constructing, validating, and comprehending proofs (and theorems) which we refer to as authentic mathematical proof activity (AMPA). To accomplish this work, we adhered to two participation related heuristics:

- *(Access) Providing access to opportunities to participate in authentic mathematical proof activity.*
- *(Engagement) Promoting participatory equity in authentic mathematical proof activity engagement.*

By access to opportunities, we mean both that students are provided tasks and prompts that may engender AMPA and attention to whether they have the appropriate tools and resources to engage in robust ways. By participatory equity, we mean whether students, regardless of background and status, are engaging in disciplinary activity in meaningful ways.

In order to support students in this activity, we identified and adapted a set of high leverage teaching practices (HLTP) studied in K-12 classrooms to the advanced proof-based setting (undergraduate classes where formal proofs are one of the primary objects of study). These practices are "routine aspects of teaching, which guide teachers to integrate students' thinking, content knowledge, and equity" (Woods & Wilhelm, 2020, p. 106). Many such practices are identified in the elementary and secondary mathematics teaching literature (Hlas & Hlas, 2012) with studies that illustrate how such practices can unfold in a classroom (e.g., Herbel-Eisenmann, 2002; Staples, 2007; Stein et al., 2008) and studies connecting the use of such practices to more equitable learning environments (e.g., Boaler & Staples, 2008).[1]We selected three teaching practices to center this contribution. These practices correspond to three components of lessons that are common to more student-centered instruction: launching tasks, managing group work, and students publicly sharing ideas in whole class discussion. We note that these are not the only HLTPs that could be designed, but we found these three to be particularly useful for planning and structuring lessons, especially in relation to our participation heuristics.

[1] We note that these practices are not always explicitly referred to as high leverage teaching practices.

In this paper, we share insights from a design-based research project focused on adopting and adapting three K-12 HLTPs to an undergraduate proof-based setting where participation was explicitly foregrounded. The project included six implementation cycles: two in a lab setting, three with a research team instructor (Author 1), and one with an external instructor. We focus on ways we engineered the HLTPs in the proof setting to better support student *access* and *engagement* in AMPA. Our contribution to literature is two-fold. First, by centering HLTPs as objects of design, we are attending to collegiate instruction in explicit ways that are often backgrounded. This is especially essential when considering how these elements of instruction may shape more or less equitable classrooms. Second, instructional design in relation to formal proof and participation is infrequently the focus of design research. The overarching design question guiding our project was: How might HLTPs be adapted and incorporated into the advanced mathematics classroom to support students in authentic proof activity? For the scope of this paper, we focus on lessons we learned during implementation of HLTPs that helped us achieve our participatory learning goals. For each HLTP, we share two instances that reflect design shifts. These shifts reflect either proof-specific adaptations that occurred when implementing the HLTP (i.e., not salient in the K-12 literature) or a substantial task refinement that occurred between implementation cycles (in service of our access and engagement heuristics).

Background on the High Leverage Teaching Practices

We operationalize high-leverage practice through an integration of Woods and Wilhelm's (2020) and Ball et al.'s (2009) definitions. Woods and Wilhem explain a high-leverage practice as, "routine aspects of teaching, which guide teachers to integrate students' thinking, content knowledge, and equity" (p. 106). Ball et al. (2009) focus on "activities of teaching that are essential to the work and that are used frequently, ones that have significant power for teachers' effectiveness." (p.461). We accordingly maintained four criteria. A high-leverage teaching practice is an aspect of teaching that (1) can be implemented routinely, (2) uses, shapes, or otherwise integrates students' mathematical thinking, (3) has potential to increase equity, access, and/or engagement, and (4) is supported by research connecting the practice to students' learning. In the section that follows, we provide evidence from the K-12 literature base that HLTPs can support students in accessing and engaging in robust mathematical activity. We note that we selected three HLTPs that meet our criteria and that also had the characteristics of being plannable and serving to help structure classroom lessons. Our focal HLTP are reflective of larger grain practices (informed by works such as Stein et al.'s (2008), practices for orchestrating discussion and Thanheiser and Melhuish's (2022) teaching routines.) HLTPs such as questioning (which is in-the-moment and not a larger structure) or lesson planning (which occurs outside of classroom instruction) were accordingly not the focus of design and refinement, though we engaged in these practices. For each HLTP, we provide a description, set of evidence, and reflection on how the HLTP has been documented to tie to equity and access. We note that the operationalizations below are a synthesis of our understanding of the practices from the K-12 literature.

Launching Complex Tasks

One essential practice is launching (and maintaining) tasks in a way that students can make sense of the task and engage productively (Woods & Wilhelm, 2020). This includes supporting students in understanding the relevant contextual features and mathematical ideas and relationships prior to engaging in problem-solving along with supporting students in developing a common language (Jackson et al., 2012, 2013). Jackson et al. (2013) documented that complex task launch was a positive predictor of students' opportunity to learn in whole class discussion. Such results are consistent with other studies of task launch and maintenance by specific instructors (Khisty & Chval, 2002; McClain & Cobb, 1998). Tools for K-12 teachers, such as Smith et al.'s (2008) task launch protocol, have served as a means to operationalize some of this work. The protocol involves reflecting on definitions, concepts, and ideas as well as what prior knowledge/relevant life experiences students need to engage with a particular task. Further, the work to make mathematical contexts, language, and ideas accessible does not have to come at the launch of the task, but also occurs in conjunction with students engaging with problems (Khisty & Chval, 2002; Livers & Bay-Williams, 2014; Moschkovich, 2013) where issues related to vocabulary or questions of context can be addressed as they emerge.

Complex task launch has been associated with increased opportunity for students of various backgrounds to learn (Khisty & Chval, 2002; Spooner et al., 2017). Mathematical tasks are embedded with contexts and content that may not be meaningful to students and can serve as barriers rather than opportunities (Sullivan et al., 2003). Attending to task launch and maintenance can serve to mitigate these issues by establishing common understanding of the mathematics and task at hand (Staples, 2007). We operationalize this HLTP as:

> The teacher engages students in making sense of tasks via attending to relevant mathematical language, ideas, relationships, task contextual features, and development of a common understanding of the task goals and context. This practice occurs prior to or in parallel with problem-solving, but does not scaffold or directly instruct on solutions to the task.

Structuring Group and Partner Work

Group and partner work often serve as an essential role in student-centered classes. Structuring these interactions shapes the opportunity for students to productively work with one another and with the mathematics (TeachingWorks, 2018). In order to do this, teachers can socially and mathematically structure group work so students have clear aims, goals, and expectations on what to do and how to interact. Webb's (2009) literature review points to variations in this HLTP including: positioning students as having diverse contributions (and describing these), providing instruction on how students can participate actively, providing explanation prompts (explicit, targeted things to talk about), focusing on questioning or debate, or role specialization. Structured group work has the potential to increase learning outcomes (Gillies, 2003) and promote high level reasoning and discussion (Cohen, 1994).

Setting up group and partner work productively is important to promote respectful interactions between students and mediate for issues of status that can lead to group work being dominated by students perceived as “high status” (Cohen, 1994; TeachingWorks, 2018). In Esmonde’s (2009) review on supporting equity in group work, they similarly point to structures such as roles or scripts that have been linked to supporting more equitable interactions where students are positioned as competent contributors of mathematics. Although, they caution that the teacher’s role includes continuing to manage groups to ensure roles or scripts are taken up. We operationalize this HLTP as:

> The teacher structures and manages partner and group work in order to engage all students meaningfully in mathematical activity. This can include scripts, clear mathematical activity expectations, and/or roles that provide guidance for how students are to interact with each other and the mathematics.

Selecting and Working with Public Records of Student Ideas

The third HLTP we consider takes place in whole class when student strategies are publicized and become the focus of discussion (TeachingWorks, 2018; Wilburne et al., 2018). Stein et al. (2008) illustrated how this HLTP unfolds as teachers anticipate, select, sequence, and then work with public records of student ideas to focus students on key mathematics. By having students present ideas and working with them publicly, common ground can develop as students have the opportunity to make sense of each other’s thinking (Staples, 2007). This teaching practice emphasizes not just students sharing strategies and ideas, but that these contributions become the focus of continued discussion. Students may be prompted to make sense of each other’s ideas, critique and debate claims and approaches (Lampert, 1990; Staples, 2007), and compare across strategies (Durkin et al., 2017). Engaging students with multiple student strategies, and in particular, focusing students on comparison, can lead to students developing more flexible mathematical knowledge.

In Jackson and Cobb (2010) reflection, they noted that discussion of student thinking plays an important role in equitable teaching. By sharing ideas publicly and engaging in discussion, students do not just have a chance to hear each other’s thinking, but “provides all students including students who are currently struggling with the particular mathematical ideas at hand, with adequate supports so that they might understand others’ explanations” (p. 5). Furthermore, by having the students present and analyze ideas, they may increase their mathematical agency (Brown, 2009), and classrooms that include public discussion of students’ multiple representations and strategies have been linked to more equitable assessment outcomes (Silver & Stein, 1996).

We operationalize this HLTP as:

The teacher orchestrates mathematical discussion where (1) students publicly present ideas and (2) students are prompted to meaningfully engage with the ideas through analyzing, critiquing, and/or comparing across student ideas.

Motivating the Study of the HLTPs in Undergraduate Proof Classes

There is a growing body of literature regarding inquiry-based mathematics education (Bouhjar et al., 2021; Larsen et al., 2013; Laursen & Rasmussen, 2019) in advanced mathematics courses, including proof-based classes. We argue there is a need to adapt and study HLTPs in proof-based classes for two reasons: (1) inquiry instruction is not equivalent to equitable instruction and thus there is a need explore ways to intentionally promote access and equity in student-centered instruction and (2) while there are substantial histories of inquiry curriculum development in proof-based courses (e.g., Larsen et al., 2013; Starbird, 2015), there is far less attention to instructional practices. Melhuish et al. (2022a) found that the majority of the literature on instruction in student-centered proof-classes focused on either student outcomes (such as student performance or affect markers) or instructor beliefs, knowledge, and instructional challenges. The instructors' role is often backgrounded in service of other research goals. For example, group work is often a substantial component of inquiry-oriented instruction (e.g., Andrews-Larson et al., 2017; Rasmussen et al., 2015); however, studies rarely address how that group work is enacted beyond description of the task. We position our study as complementary to this literature on student activity and curriculum, but unique in that we expand the object of design to incorporate specific elements of teaching, namely our three focal HLTPs.

Finally, we recognize the need to study the HLTP in these courses rather than just directly adopt them from the K-12 setting. While we hypothesized the fundamental roles of the HLTPs may stay consistent, their enactment in proof-based courses is likely to be shaped by the unique context of working in the formal representation system of proof (Weber & Alcock, 2004). There is a substantially increased level of abstraction (Hazzan, 1999) and new ways of argumentation that are beholden to idiosyncratic mathematical conventions (Lew & Mejía-Ramos, 2019) and specific norms and values of the mathematician community (Dawkins & Weber, 2017). Much of the literature in this area points to the challenge of this transition for students (Stylianides et al., 2017), and thus we anticipated that engaging students in authentic mathematical proving activity will be a non-trivial task and involve substantial intentionality and engineering.

Authentic Mathematical Proof Activity and Participation Heuristics

In undergraduate proof-based courses, typically taken by mathematics majors and future secondary mathematics teachers, the primary object of study becomes the formal mathematical proof. We take a participatory stance on student learning borrowing the notion of *productive disciplinary engagement* from science education (Engle & Conant, 2002). That is, our goal was to engage students in activity that resembles the work of research mathematicians. We hypothesized that HLTPs can support our student activity goals by providing structures and mechanisms to engage students meaningfully with tasks and each other.

Proof Foci of Tasks and Activities: Proof Construction, Validation, and Comprehension

Most of the extant research in undergraduate proof settings focuses on students' abilities in the realms of proof construction, proof validation, or proof comprehension (Selden & Selden, 2017; Stylianides et al., 2017). Proof construction can be broadly conceived of as the development of an argument which contain conclusions (the statement to be proved), data (which provides the foundation of the argument), and warrants (which provide the justification to connect the data to the conclusion) in alignment with Toulmin (1958) argumentation scheme (Simpson, 2015). A mathematical argument is then a formal proof when it "dr[aws] on symbolic notation and logical reasoning" (Fukawa-Connelly, 2012, p. 333). The proof construction process can stem from informal ideas such as those that come from exploring examples or diagrams that can then be formalized through activities such as elaborating, syntactifying, and rewarranting (Zazkis et al., 2016).

While proof construction is most prevalent in the literature (Mejía-Ramos & Inglis, 2009), proof validation is also an important aspect of mathematician activity (Weber, 2008; Weber & Mejía-Ramos, 2011). Weber and Alcock (2005) have suggested validating a proof is exploring whether "If (a subset of the previous assertions in the proof), then (new assertion)" (p. 37) is warranted at each line of proof. Studies have suggested that mathematicians validate in two phases: determining the structure of the argument and then checking each line of the argument. As such, validating activity may be identified through the lens of organizing information into what is known and what needs justification, evaluation of warrants of claims, and appropriateness of proof structure.

Finally, proof comprehension is an essential aspect of mathematicians' activity (e.g., Melhuish et al., 2022b; Weber & Mejía-Ramos, 2011). Mejía-Ramos et al. (2012) have developed a framework for assessing proof comprehension highlighting two main dimensions: local understanding (which can be gleaned from a small number of statements within a proof) and holistic understanding (which cannot). In particular, their model identifies three aspects of local understanding – meaning of terms and statements, logical status of statements and proof framework, and justification of claims – and four aspects of holistic understanding – summarizing via high-level ideas, identifying the modular structure, transferring the general ideas or methods to another context, and illustrating with examples. Thus, we operationalize proof comprehension as attending to local and global aspects of an existing proof to understand both the argument and its constituent parts.

Authentic Mathematical Proving Activities (AMPA) and Participation Heuristics

In order to account for student activity in classroom settings, we developed the Authentic Mathematical Proving Activities (AMPA) framework (Melhuish et al., 2022b) via a synthesis of the literature on mathematician activity to further operationalize their objectives (objects, motives) and tools in their activity systems. The *objects* of activity include proofs, concepts, and propositions. With regards to each of these objects, we identified three main *motives* related to these objects reflective of comprehending, constructing, and validating. Activity then consists of *objectives,* combining these objects and motives (e.g., constructing a proof, comprehending a proposition). *Tools* are then used to achieve these motives. For the purpose of this paper, we do not expand on all of the framework's tools, but provide some examples to situate our goals and results which focus on instructional elements more so than student activity. We note that these tools include processes such as *analyzing and refining* (the activity of exploring and modifying an extant object (proof, statement, or concept) via examining assumptions and implications) or *warranting* (identifying implicit/explicit warrants in a particular claim). They also include other resources such as using *diagrams*, *examples*, or *logic.* We consider a student engaged in AMPA when they are taken on authority and agency in using disciplinary tools towards disciplinary objectives.

With this goal in mind, we elaborate the two participation heuristics we shared in the introduction:

- *(Access) Providing access to opportunities to participate in authentic mathematical proof activity.*
- *(Engagement) Promoting participatory equity in authentic mathematical proof activity engagement.*

The access-related heuristic focuses on whether students have the opportunity to engage in tasks that can lead to AMPA. Such opportunities depend both upon the tasks teachers provide and whether students have the necessary resources and understandings to engage in intended ways. Access is infrequently uniform across students and those students who more quickly draw on definitions, theorems, relevant understandings, and accurate interpretation of formal mathematical language may have increased access to AMPA (e.g., Moore, 1994; Weber, 2001; Weber & Melhuish, 2022). Thus, a driving feature of our design is maximizing access to opportunity for AMPA. The engagement heuristic helps us attend to whether students realize these opportunities in equitable ways. That is, are all students taking on the disciplinary activity in meaningful ways? Brown (2018) and Reinholz et al. (2022) have documented ways that students (particularly those of minoritized backgrounds) may not have equal opportunity to participate in small groups and whole class discussion, respectively. As argued by Brown (2018) and Johnson et al. (2020), inquiry is not a panacea for equitable instruction, and equitable instruction involves intentionality in instructional practice beyond providing rich tasks. We see both heuristics as essential to our instructional engineering of high-leverage teaching practices.

We provide clarifications regarding these foci. First, they are not independent. Without access to opportunities and resources, equity in participation cannot occur. Second, we are not attending to a number of aspects of supporting equitable learning environments that go beyond our central focus on participation in AMPA. We take caution not to overstate claims of creating equitable classrooms. Finally, we note that initial design focused on the first heuristic; however, observations about disparities in participation and in what ways during early cycles led to the explication and attention to the second heuristic with intentional modifications in later cycles.

Methods

The data from this paper comes from several cycles in a design-based research project (Cobb et al., 2003). The project focused on the development and refinement of HLTPs in the context of three introductory abstract algebra lessons. We are using a design-based research approach due to the project aims of theorizing and developing curricular materials. As Cobb et al. elaborated, design-based research contains five features. First, the research involves developing "theories about both the process of learning and the means that are designed to support that learning" (p. 10). For our project, we take a participatory lens on learning placing HLTPs and their relation to student engagement in AMPA as the focus of theorizing. The second feature is that our project is highly interventionist. We are studying instruction and learning as it plays out. We note, that means, "the study of phenomena as complex as learning ecologies precludes complete specification of everything that happens" (ibid., p. 10). Unlike experimental research, we are not attempting to account for all variables, but rather are studying a system with forefronted planned elements (HLTPs, tasks) while other elements are backgrounded. The third feature reflects the prospective and reflective nature of these experiments. We came into our work with a theory of how HLTPs and specific proof tasks may support student engagement in participation. As the cycles of the experiment played out, we developed more local conjectures and detailed understanding of mechanisms involved using many levels of analysis. This leads to the fourth feature, the iterative design of this type of research where conjectures and evidence lead to revision. The final feature reflects the nature of the theories produced. They are not global learning theories, but rather local to the problem targeted by the design experiment. At the same time, the insights developed should not be so constrained to a particular setting that others cannot make use of the insights. In order to meet the final features, we engaged in many cycles of theorizing and design, implementation, analysis, and refinement in different settings to increase transferability. Presenting our findings as lessons learned and shifts made in iterative design represents our attempt to report in a manner that is true to design research and likely to render our specific insights adaptable to other researchers and instructors.

The Focal Lessons

Each of the lessons was designed with the primary goal of engaging students in validating, constructing, and comprehending proof, respectively. The lessons were designed to take one class period of 1 hour and 20 minutes; however, depending on implementation some lead-up or wrap-up work occurred in the class session before or after. Lesson 1 (which we refer to as the Structural Property Task) focuses on the theorem: *Let G and H be isomorphic groups. If G is abelian, then H is abelian.* This theorem was selected because it is a common type of theorem in abstract algebra and students often approach it in two different ways (Melhuish et al., 2019): 1) beginning with elements in G and showing their images commute or 2) beginning with elements in H and showing they commute. This allows for students to investigate the differences between approaches, *validate* the approaches, and refine the proof or alter the statement to only use the necessary assumptions. Lesson 2 focuses on *Lagrange's Theorem.* This theorem was selected because the key idea can be apprehend via example exploration (e.g., Leron & Zaslavsky, 2013). Students investigate example groups and their cosets, attend to the multiplicative structure (key idea), and develop a set of lemmas about cosets to *construct* the proof. Lesson 3 focuses on the *First Isomorphism Theorem* which was selected due to the complexity involved in the proof (Nardi, 2000) and has students engage in *comprehending* the statement (via example exploration) and the proof (via identifying structure and line-by-line explorations). See Appendix for a complete outline of the final version of the lessons.

Design Cycles and Setting

Participants were undergraduate mathematics majors (some dually earning high school teaching certification) from a large, research university in the United States. The participants had all completed a transition to proof class. Students from the lab setting had completed abstract algebra and students in the classroom were several weeks in and had been exposed to basic definitions and proof techniques. Information from each cycle can be found in Table 1.

The design process entailed several phases of developing and modifying focal tasks, planning enactment of HLTPs, and hypothesizing corresponding student activity in terms of participation in AMPA. The tasks were implemented, first in a lab setting, then in the classroom. The lab settings (where students engage with tasks and instructor-researchers outside the confines of a classroom and full-sized class) were led by the first author as instructor-researcher and third author who observed and interjected questions and prompts at relevant times. Author 1 implemented the first three classroom implementations. The classroom implementations were interrupted by pandemic protocols ending cycle 3 prematurely (after lesson 1) and leading to cycle 4 occurring online. We conducted the first two cycles in a smaller setting to better allow for testing out the relationship between our tasks, HLTPs, and student activity. This afforded greater attention to student thinking and participation and removed some constraints of the normal classroom. Through the two lab setting

Table 1 Overview of six cycles and student demographics

	Term[1]	Lessons	Modality	Instructor	n	Demographics
Cycle 1	Spring 2019	All 3	Lab Setting	Researcher	3	3 Women
Cycle 2	Summer 2019	All 3	Lab Setting	Researcher	4	3 Men, 1 Woman
Cycle 3	Spring 2020	Lesson 1	Classroom	Researcher	15	6 White Women, 1 Multi-Racial Woman, 4 White Men, 4 Hispanic Men
Cycle 4	Fall 2020	All 3	Online Classroom	Researcher	29	7 White Women, 3 Hispanic Women, 4 African-American Women, 7 White Men, 5 Hispanic Men, 3 Asian Men
Cycle 5	Fall 2021	All 3	Classroom	Researcher	17	2 White Women, 1 Multi-Racial Woman, 3 Hispanic Women, 6 White Men, 4 Hispanic Men, 1 Man (unknown ethnicity)
Cycle 6	Spring 2022	All 3	Classroom	External Instructor	13	2 White Women, 1 Hispanic Woman, 3 White Men, 6 Hispanic Men, 1 Asian Man

[1]Tasks occurred in the first two thirds of a term allowing for several months for analysis and modification between cycles

implementations, we developed more precise conjectures about the relationship between tasks, HLTPs, and student activity and made modifications for the classroom. We initially planned only two classroom implementations. Since Cycle 3 and Cycle 4 were affected by pandemic protocols, we added Cycles 5 and 6, which were fully in-person. During the final implementation, we shifted from a research team instructor to an outside instructor to extend our movement toward less controlled and more naturalistic instructional settings. This allowed for the tasks and implementation guide to stand on their own and not be shaped by unaccounted knowledge that was linked to being part of the design team. Additionally, an expert review panel provided feedback at key points (before cycles 1 and 3).

For each cycle, the task implementations were video-recorded and transcribed. Research team members (including Author 2 and 3) observed and took field notes on implementations attending to ways that students did or did not engage in hypothesized activity including attention to equity in this participation within small groups. We engaged in both "design minicycles" and retrospective analysis in accordance with Cobb and Gravemeijer (2014). Design mini-cycles include debriefs after each lesson with the project team to come to consensus about the ways in which the lesson enactment (including HLTPs) aligned with hypotheses in terms of supporting students in AMPA related to the lesson goals. In between each cycle, the project team met to reflect more holistically on the prior implementations, revisit important points in the data, and, in some cases, conduct extended analysis of particularly salient moments. We used a number of analytic tools in these various stages of analysis, including:

- the AMPA framework (Melhuish et al., 2022b) on data from cycles 1–3,
- The Math Habits Framework (Melhuish et al., 2020) to analyze instructional moves in our initial plan and after cycle 3 implementation (the first in a classroom), and
- The Activity, Authorship, and Animation (AAA) Authority Framework (Hicks et al., 2021) on data from cycle 2.

We selected this set of frameworks in order to gain more systematic insight into how our conjectured instructional approach and implementation linked to student activity, both in terms of the access heuristic (operationalized via AMPA) and the engagement heuristic (operationalized via analysis of who contributed to the AAA components). The results from these more systematic analyses complemented our broader design minicycle analyses. We revised our plan for implementation and/or refined our use of the HLTPs whenever we identified evidence of disconnects in access to or engagement in proof activity, such as noticing students did not appear to have needed tools to make progress on an activity (an access issue) or that certain students dominated conversation (an engagement issue). Additionally, we identified spontaneous elements of instruction (e.g., instructor prompts that helped students past an impasse) and student activity (e.g., features of students discussion that led to productive communication) that appeared supportive to our access and engagement aims, and often incorporated these into our task materials.

We began with rather primitive hypotheses linking HLTPs to supporting students in rich discussion and participation in proof activity. Initially, we designed each lesson to have a focal HLTP and proof activity pairing: selecting and working with public records of student ideas and proof validation, using and connecting mathematical representations and proof construction, and launching complex tasks and proof comprehension. We paired the practices intentionally drawing on (1) the idea that students may be focused on important structural aspects of their proofs via comparing public records (and thus attend to validating them), (2) working with visual records can support apprehending key ideas needed to construct proofs, and (3) complex task launch to support accessing and comprehending ideas in proofs. However, we quickly discovered that trying to isolate practices in this way oversimplified a complex setting and relegated the role of non-focal HLTPs to the background. During later cycles, we attended to all focal HLTPs throughout each lesson and in relation to different proof activities, although the hypothesized pairings remained some of the most salient. Additionally, we more thoroughly incorporated the HLTP of Structuring Group Work midway through the project when our analyses pointed to unbalanced participation in group work. We note that while visual records remained a part of our design, we background this HLTP this manuscript as it has a more thorough treatment in undergraduate settings.

At the completion of the data gathering, we engaged in a retrospective analysis focused on the characteristics of the HLTPs as mechanisms to support students in engaging in AMPA. The crux of our design focused on hypotheses related to ways that carefully planned and designed HLTPs (in conjunctions with tasks) can support students in access and engagement in AMPA as they comprehend, construct, and validate proofs. This analysis differs from the cognitive analyses most common in studies of proof-based learning contexts. We are not attempting to make claims about students' knowledge or evolution of knowledge. Rather, we focus our claims specifically on the links between HLTPs and participation. Thus, we leverage our data corpus, including the various analyses described above, to provide images of how the HLTPs supported access and engagement in AMPA. The theory arising from our design experiment also entails aspects of student understanding of particular topics and their learning about abstract algebra, but our goal in this paper is to portray aspects of the emerging theory relating instructional moves and participation in proof activity. The stories and challenges described in the results section convey the lessons we learned that constitute a core of that part of our theory. We share examples from different points in our design because they were the occasions that led to modification and elaboration of our understanding of how HLTPs can support AMPA. The examples span the duration of the project, and in each case, we share evidence of what played out uniquely in the proof context (thus, elaborating our understanding of HLTPs in the proof setting), and for instance that led to modification, we provide some detail of the impact of particular changes.

Results: Lessons Learned from the Orchestrating Discussions Around Proof Project

In order to share insights from this design project, we present our results as a series of episodes and noticings that occurred through the design and instructional engineering process. We organize these results by situating these instances within each HLTP to which they were most related. However, this treatment is done for readability purposes rather than implying that each teaching practice operates disjointly. In fact, teaching practices can and often do overlap. For example, students may engage in structured group work around a public record from discussion. Further, ramifications of decisions made in service of one HLTP can influence activity in many elements of a lesson.

Launching Complex Tasks in the Proof-Based Setting

In the K-12 setting, launching complex tasks involves making sense of task contexts, questions, and anticipating key mathematical structures and relationships that might support students in productive problem-solving (adapted to proving activity for our work). We designed task launch to include (1) unpacking key definitions and relationships in theorems and (2) using examples and visual representations to make sense of key ideas in theorems. These two activities were designed to anticipate proof structure (Samkoff & Weber, 2015) such as providing a definition that anticipates a proof structure (e.g., definition of abelian when showing a group is abelian) or provide insight into key ideas (Raman, 2003) and structural features (e.g., seeing the structure of cosets to structure the proof for Lagrange's Theorem). Further, we keep a public record of definitions and other key ideas students may need to provide a resource for students' mathematical activity to build from.

We argue that these are essential elements needed for productive proof engagement based on the multitude of literature that suggests novice provers understanding of concepts and definitions (e.g., Moore, 1994) greatly shapes their proving activity and that identifying and working from key ideas is more consistently found in expert provers' practice (e.g., Raman, 2003). In this section, we share two learnings from implementing complex tasks launches. First, we discuss teaching prompts related to object references and quantification, a type teacher prompt that was important in proof setting, but not emphasized in the existing K-12 literature that informed our design. Second, we share a major modification we made as a result of finding that students were not anticipating proof structures in the ways we initially hypothesized for the proof construction task (Lagrange's Theorem.)

Increasing Support of Students' Access to Formal Mathematics Through Identifying of Mathematical Object Referents and Quantification

While many of the instructional prompts we documented directly paralleled those found in the K-12 literature, *mathematical object referents* and *quantification* seemed unique to the undergraduate setting. Furthermore, the need to support students in

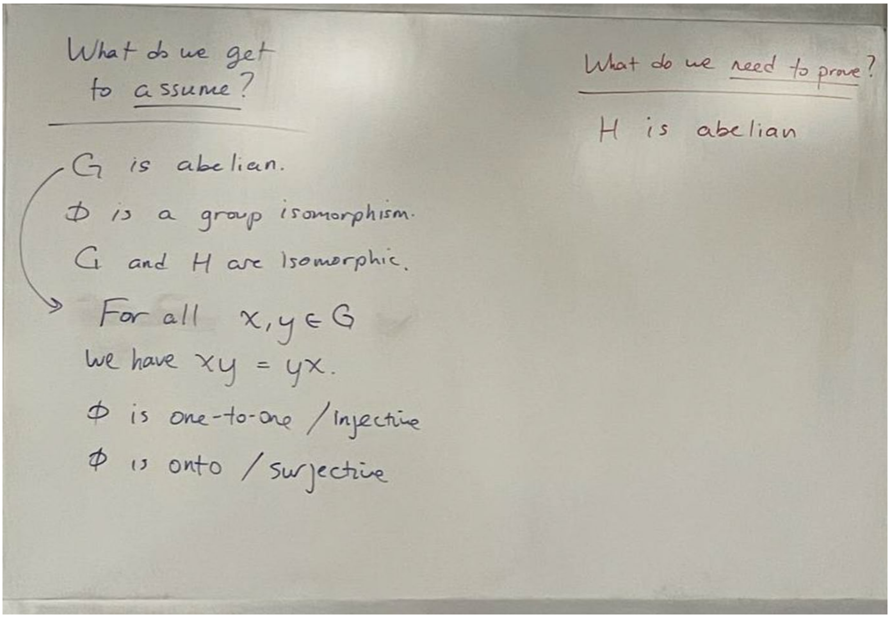

Fig. 1 Public record of student assumptions, definitions, and conclusions for the structural property task

recognizing *mathematical object referents* and the role of *quantification* occurred across implementations and types of lessons, and ultimately became a planned part of instruction in later rounds. Because these tools (such as definitions, statements, and their referents) were needed to engage students in the different types of activities to come, we chose to give class time to unpacking these during complex task launches to better support students in developing shared understandings that would carry through the lesson. We use the structural property task as an illustration. As students suggested the assumptions and conclusions, common responses include "one-to-one" and "onto" without referring to the mathematical object that has these properties. In one implementation, a student offered "G and H are a group isomorphism." This led the instructor to ask "who do we call isomorphism?" to which another student responded, "G goes to H" which again the instructor asked "What do we call that?" with the student responding, "ϕ". Figure 1 represents the public record of student suggested ideas and definitions. Similar conversations have occurred about asking what type of "object ϕ is?" to draw attention to ϕ as a function and asking whether a "homomorphism" was a function or a property. This relates to the subtle issue of properties being defined by the existence of a function. Furthermore, during these exchanges students tended to provide unquantified statements. When asked to unpack abelian, a student suggested "$xy=yx$" in one class and "There's an a and b in the group that also, a operated with b also equals b operated with a" in another class. In both cases, the quantification remained unclear and the instructor followed up to ask questions such as "is it for some a, for all a, are these arbitrary?" with many students in the class clarifying, "for all a."

If we reflect on our guiding heuristics, these two types of prompts were recognized and intentionally integrated into future iterations for several reasons. First, quantification plays a key role in productively engaging in the proving and validating activity, both generally and in our tasks. In some sense, the difference between the "valid" and "invalid" proof approach to the structural property task is attention to the role of "for all" in the abelian definition (and how it gets proved using arbitrary elements from *H*). Not attending to the role of "for all" in their proof comprehending activity limited their opportunity to recognize the proofs' validity. Second, discussions about mathematical objects and their referents became vital when considering access and promoting opportunities for all students to engage. While some students were immediately able to engage in constructing, comprehending, and validating proofs with complex levels of objects and symbols, other students would hit an impasse at different stages. For example, when constructing initial proof approaches in the structural property task, some students were unable to get started as they lacked the necessary tool of introducing a ϕ to build an argument. The importance of referent objects became a significant compounding factor in later activity such as dealing with the FIT where the presence of multiple functions can lead to the proof becoming impenetrable (see Nardi, 2000). In our early lab settings and online implementation, we found that without explicit attention to symbols and referent objects, students made little headway into mechanics of the proof. This is an idea we will revisit in the group work section.

Supporting Students' Engagement in Proof Production and Proof Understanding Via Unpacking Structural Elements and Meaning Beyond Formal Definitions

In our initial implementation, we focused primarily on objects and formal definitions as part of complex task launch. This was often accompanied with exploring a few examples to notice structural features and tie features of the examples to their role in the focal theorems. However, we found that such exploration may remain disjointed from future activity without active anticipation and focus on how a structural noticing may carry over to a proof context. This was particularly apparent in the Lagrange Theorem task that hinged on students noticing the key idea that cosets induce a multiplicative structure on the elements in a group. We initially hypothesized that students could produce a multiplication argument by unpacking the statement and arriving at a multiplication goal (WTS: $|G|=k|H|$ for some k) and then connecting $|G|$, k, and $|H|$ in their diagrams. This link was tenuous for students. We made the most substantial modification after the first implementation of the Lagrange Theorem task. We will briefly share how the first implementation played out and then data from the second implementation that was more productive.

In the first implementation, the students had explored multiple examples of groups and subgroups and identified where they could see the parts of the theorem, but hit an impasse as they attempted to use their examples to build their argument focusing on showing there is "no remainder." Anna suggested, "What if we did a couple of cases, like where the order of the group was even, or it was odd? If it's even, you have … Then it's just $2\,k$ and if it's odd, $2\,k+1$." Elena continued this line of thinking addressing various cases reflecting different "factor[s]" and the students elaborated that they needed to show there would be no remainders. However, after some work, that last of the trio, Elsa commented, “It would be a really big proof.” First, we want to note that the students’ approach was quite reasonable. We conjecture that they were relying on prior proof experience where number theory arguments about division often rely on particular cases. They focused on not having a remainder. That is, they were drawing on prior strategic proof knowledge rather than drawing on coset explorations to formalize. Ultimately, the instructor heavily scaffolded the connection and, we would suggest, was the only one engaging in AMPA by the end of the lesson. This did not fulfill our overarching goal to engage students in AMPA where informal activity (example-based) and formal activity (proof) served a mutually supporting role (what some researchers may suggest to be cognitive unity, Garuti et al., 1998).

As a result of this experience, we hypothesized that additional instructional support may be needed to help students draw upon their informal exploration in formal contexts. In this particular case, we expanded the task launch to include not just formal definitions, but also having students recall more informal ideas about multiplication that can serve to bridge between the activities. As students considered their multiplicative statement “WTS: $|G| = k|H|$ for some k”, the instructor-researcher prompted them to “[c]aptur[e] this with a visual. How is this illustrating what we mean by multiplication?” They also provided specific numbers $12 = 3 \times 4$ prompting “let’s think back to elementary school when we write these things, and we’re gonna make a similar type of visual to go with this that’s kind of connected to what we mean by multiplication. See if you can also sketch something out that goes with this idea of $12 = 3 \times 4$.” After some partner discussion, the instructor-researcher then guided a full group discussion about a definition for multiplication that built from student suggestions of “repeated addition” and “totaling up.”

After some additional exploration and lemma generation about cosets, the instructor-researcher prompted the students in this group in much the same manner, “So if these three lemmas are true, how might they help us establish the multiplication structure that we were trying to get up here?” In this case, the students were able to translate between informal and formal with one student, Jasmine, explaining, “Oh, so like the union of the cosets is G is basically k x H.” Asked to repeat, she elaborated, “The first lemma, that the union of the cosets is G. The repeated addition kH. I mean, when you merge them together, you get a G.” The conversation continued with the students connecting each of the lemmas to their role in the multiplicative structure. Notice that the students are drawing on the shared language of “repeated addition.” We saw this as evidence that work done at the task launch supported the students in engaging in more authentic proof activity.

This example is emblematic of a larger activity trend. We observed that the switch to formal proving often primed students to draw on formal proof knowledge to the exclusion of informal explorations. The proof construction task hinged on leveraging the “key idea” of multiplicative structure, and thus needed intentional engineering to engage students in using their informal understanding of multiplication in relation to the formal proof. We suggest this result generalizes as key ideas are by definition a means of connecting informal and formal. Such a connection may be obvious to a more experienced prover; however, it needs explicit parallelism for a more novice prover to use their informal ideas to support proof construction. Throughout different tasks and HLTPs, there was a need for instructors to orient student ideas such that there was consistency and connection across informal and formal representations. We return to this theme in other sections.

Reflection on Launching Complex Tasks in the Proof Based Setting

In many ways, we were able to import the primary essence of complex task launch from K-12 mathematical settings. Our implementations suggested several nuances that are likely proof-context specific (or at least more salient in this setting.) These considerations were primarily access-driven. First, attention to mathematical objects and their referents is crucial to developing a shared language and providing the basic tools for students to engage in activity. Second, quantification is a huge aspect of definitions and particularly how definitions relate to proof structures. Students' descriptions of mathematical ideas may lack that level of precision – and for good reason. That level of precision did not serve much purpose in non-proof based classes. However, in these contexts it is essential and can support later activity. Third, one of the most challenging aspects of complex task launch was anticipating ways to support students in not just seeing important structural relationships (which is an element of this work in other mathematical settings), but the tools needed to link structural relationships in an informal discussion to later formal proof activity. We suggest explicit attention to ideas that may bridge and anticipate proof structure (beyond just formal definitions) that can serve to alert students in making connections.

Structuring and Managing Group work

A key component of group work is designing and developing tasks that are groupworthy (Lotan, 2003). In the context of design-work in proof-based classes, the focus is often on the nature of the task, a necessary component for group work where students may work on challenging proof construction or a task trajectory that supports reinvention of formal mathematical concepts (e.g., Larsen, 2013). However, in our engineering we also attended to instructional choices about the structuring and management of this work–how would students actually do this work in a group setting? Initially, we relied on two mechanisms for structuring group work, "think-pair-share" (Kaddoura, 2013) and partner exchanges (similar to peer review, Reinholz & Pilgrim, 2021). Such approaches have been documented in proof-based instruction; however, in accordance with our guiding heuristics, we found a need for more complex and intentional structures to promote more equitable participation in AMPA. In terms of group work, we identified this need from several perspectives. In the initial in-class implementation, we witnessed imbalance in how students participated with their partners. A later analysis of the lab setting experiments also pointed to inequities in participation (see Hicks et al., 2021). Finally, during the online implementations, we noted that unstructured group time often resulted in both low participation by some members of groups and, in some cases, no student activity met our definition for AMPA. In this section, we share two types of modifications we introduced to group work between cycle 4 (the online cycle) and cycle 5 (the first full in-person classroom implementation.) Both draw on ideas from complex instruction to promote more equitable group work (Cohen & Lotan, 1997) including expanding expertise (a more thorough treatment of this idea can be found in Weber & Melhuish, 2022) and distributing responsibility. We begin by sharing data from the structural property task that contained a partner exchange to show how this type of structure may be insufficient to promote equitable participation across partners. We then share data from the FIT task where a think-pair-share structure was also initially insufficient and ways that we incorporated more intentional sharing of responsibility amongst group members.

Adjusting Partner-Exchange Structures to Increase Participation and Decrease Status Disparity During Proof Comprehension Phases

Differences in students' comfort in beginning the production of a formal proof is one source of inequitable participation in this context. Apprenticeship into formal proving requires a fundamental shift in argumentation and language. Knowing "where to start" is a substantial hurdle for novice provers who are only beginning to develop strategic knowledge (Weber, 2001) for operation in this system. As such, a status imbalance can occur between students who are comfortable with formal proofs and those less so. To illustrate this issue, we turn to the Structural Property Task. Our initial design involved students working with partners and exchanging their proof approaches. This was structured such that students were instructed:

> I want you to come up with one thing that makes sense about what [your partner] did and one thing that maybe you have a question about in regards to what was playing out or, how things are labeled, anything you can have a question about in here. Pull out your approach, exchange it with your partner, spend about two minutes reading through it, seeing if you can come up with one question and one thing that makes sense.

The instructor guided students in exchanging and taking on these roles. However, closer inspection of this activity revealed that these structures did not always play out and that certain students took on an "expert" role while their partners did not. Consider the following exchange:

> Aiden: so I ask about your [pause].
> Brianna: I guess. Even though I don't know anything.
> Aiden: I think there's a problem with -- so you say, "since *G* and *H* are isomorphic, *G* and *H* are 1-1, onto, and homeomorphic"
> Brianna: Oh, I was referring to um the property that she gave us, and then.
> Aiden: but-- yeah I think the problem is that it's just missing-- the thing that's 1-1 and onto is the function between them so it-- I don't know if it's right to say that G and H *are* 1-1 and onto but I would probably say there exists a function from G to H that is 1-1 and onto

Aiden then guided Brianna in the construction of a new proof.

We conjecture this disparity occurred because Aiden had a mostly complete proof from the night before whereas Brianna had a set of initial ideas. We had moved the proof production portion to an at-home activity to not impose time constraints on the initial proof construction (an access decision); however, this did not mitigate the issue as only a handful of high status students (students who frequently participated at high rates) brought mostly complete proofs.

In the next iterations of this lesson, we transitioned from students discussing their own proofs to ones provided to them. These proofs were student-generated from a prior study (see Melhuish et al., 2019). Each student in the class was given either Proof A (an argument beginning with arbitrary elements of the domain) or Proof B (an argument beginning with aribitrary elements of the co-domain) and were provided private reasoning time to make sense of the proof in front of them with the instructor explaining [Cycle 5], "You're kind of now the *expert* on, on the one in front of you. So I'm going to give you a couple minutes to try to digest it and think about, 'can you explain what's going on in this argument to somebody who doesn't see it?'".

This approach led to robust conversations where we did not observe the same sort of status disparities or divergence from the intended activity. For example, consider the following partner discussion with Isabella (Proof A) and Jake (Proof B). Isabella explained her proof, "I'm going to say that this is ϕ. So it's letting a,b be elements of G, so there exists a ϕ that a operation b is equal to $\phi(a)$ times $\phi(b)$ […] which also if you have $\phi(ba)$ is equal to $\phi(b)\phi(a)$ since G is abelian." Her partner then comments on what makes sense and asks a question about the connection to the codomain group H. The partners work together to summarize the main idea:

Jake: They're just trying to show that…
Isabella: But they are showing that either way you write it…
Jake: That the [inaudible] no matter which way you would put it, would be okay.
Isabella: That's why it is written three different ways, to show that, no matter which way, they're all equal.

The partners then exchanged roles with Jake leading a discussion of the other proof approach and Isabella commenting on features of the proof. If we compare this conversation to the previous, we can see that both students are engaged in what we would call AMPA reflected in using tools, like summarizing, to engage in proof comprehension.

By switching the focal comprehension object to existing proofs, both students were positioned to have expertise (on their respective proofs) and there were more entry points into the activity. A natural critique of this modification is that the proofs are no longer stemming from students in the class. However, we highlight that this move opened an avenue for additional competencies and provided a means to support comprehension activity in more equitable ways. We share this example for a couple of reasons. First, this type of "construct a proof task" is very common to proof courses with active student engagement. Yet, there are substantial differences in students' comfort, access, and knowledge of the formal proof construction process. This may sometimes lead to quite different classroom experiences for different students (for example see Dawkins et al., 2019). If other designers share access and engagement goals, it is worth being strategic about when and how proof construction tasks are used. Unguided open prompts to prove may inadvertently amplify status differentials.

Increasing Student Authentic Activity and Student Participation by Delegating Responsibilities to Engage with Formal and Informal Mathematics

In the lab setting (cycle 1 and 2), the instructor-researcher often asked students a series of targeted questions when they encountered challenges moving between formal and informal systems. In the full classroom context, an instructor no longer has the ability to engage in conversation with all students in small groups. After the online implementation (cycle 4), we found that when students encountered such challenges, they often did not have the tools to move beyond an impasse. Prior to cycle 5, we developed more sophisticated group structures (rather than just think-pair-share) in order to engage students in more authentic activity and to assure more voices were heard. Many of these modifications were in service of the FIT task as the formal proof involved making sense of abstract and layered arguments working to move the responsibility from instructor-researcher to students to orchestrate group discussion.

One strategy that proved useful was converting instructor prompts that were fruitful into questions for students to lead discussion about. Many of these prompts shared common features with discussion elements from complex task launch. To illustrate, in the FIT task during cycles 5 and 6, the classes were subdivided into four groups and each group was given one section of the FIT proof to be responsible for explaining to the class. In order to promote more equitable participation, each member of the group was given one question and tasked with leading the discussion on that question. This provided support to engage in deeper AMPA and a mechanism to engage all students in having a meaningful role. The instructor launched the activity stating, "So, if you're person one, your job is to bring this question to everybody and make sure you talk about it and resolve it as a group. But you definitely don't have to do them individually." They continued to clarify that the "questions build on each other" so students could not just work on their questions independently. Notice the focus is on both the responsibility of the individual but the need for collaboration amongst the group.

The questions were derived from using Mejía-Ramos et al. (2012) proof comprehension framework in combination with key referent object prompts from earlier implementations. For example, each group had one student lead a discussion on, "What is the difference between β and ϕ?" in terms of the domain and codomain elements (the isomorphism and homomorphism maps, respectively). Across both implementations using this mechanism, all members of the small groups engaged in the conversations and a member of each group was able to come up to the board and provide rather sophisticated explanations of their proof sections addressing the referent objects accurately and warranting lines in the proof. We used roles and responsibilities for both constructing new objects (see Incorporating Public Records of Partial Information to Promote Access to Formal Ideas in Relation to Students' Informal Ideas), and in proof comprehension activities (including the FIT and partner exchanges from the structural tasks.) As this was a later adaptation to our sequence, we have less evidence of how structuring group work in such manners could occur throughout the tasks. However, we conjecture that this is a transferable mechanism. That is, using roles and responsibilities that are mathematically meaningful (such as leading a discussion around important aspects of a proof or example), tied to proof activity aims (such as comprehending), can serve to support more equitable participation in terms of all students contributing to the group work. Specific to the proof setting, we aimed to incorporate roles and responsibilities that can serve to navigate between formal and informal.

Reflection on Structuring and Managing Group Work in the Proof Based Setting: Two Shifts We Made

One of the more challenging aspects of working in the proof-based setting is that the abstract nature of the mathematical content and the formal ways of arguing and communicating often privilege a particular set of competencies that may exacerbate status issues. In the literature, this is sometimes approached via differentiation of instruction such as having easier and harder proofs for students to engage with (Dawkins et al., 2019). While such a mechanism may work in a more traditional IBL setting with a high degree of independent work, it does not easily import to a more collaborative setting. We found that we needed to evoke a range of competencies (such as comprehending, validating, explaining a proof, generating and analyzing examples, rather than just constructing proof). We also found we needed to carefully delegate responsibilities such that there were multiple ways to remain involved. This was often partnered with assigning students tools to open opportunities for authentic activity such as guiding questions or particular structures for example creation. While in some ways, this scaffolds the activity further than our initial design, we were able to document more AMPA amongst more students than in our less structured group work attempts.

Selecting and Working with Public Records of Student Ideas

Two types of public records drove this practice across our lessons: records of examples and records of proofs and proof elaborations. By proof elaborations, we mean students may or may not have developed the initial proof version, but publicly share their understanding of the proofs through recreation or elaboration (e.g., identifying objects and creating small deductive subproofs when needed). As in the K-12 setting, comparison was intended to highlight structural attributes. In the proof setting, this may highlight structural differences as between proof frameworks (Selden & Selden, 1995) or structural commonalities that may anticipate a proof, such as the examples explored in service of the FIT and Lagrange's Theorem discussed in the prior sections.

Attending to Whose Proofs and Products are the Focus of Engagement and Moves to Incorporate More Opportunity for Less Vocal Students' Ideas to be Publicly Discussed

A key component of this work is anticipating the ideas that may be selected, sequenced, and compared (Stein et al., 2008) For example, the two proof approaches

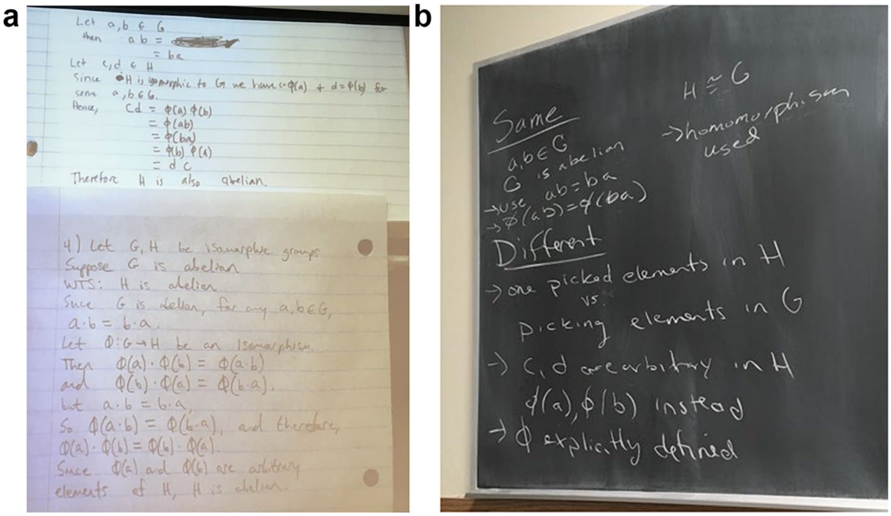

Fig. 2 **a** The two student approaches to proving the structural property theorem. **b** Public record of students identified similarities and differences across the approaches

structural property task stemmed from the most common approaches by students in survey study (Melhuish et al., 2019). However, we found that pre-planning which proof strategies to select for comparison limited whose ideas would make it into the public space. In our first classroom implementation (cycle 3) relying on this approach led to the selection of two, vocal, white mens' proof approaches becoming the focus of the conversation for the duration of the class. The two students, in sequence, shared their proof approaches. These two approaches were anticipated and selected because they began the proofs in different locations (starting with elements from G and from H) and contrasting the proofs can draw attention to important structural features. After both students presented their proofs (see Fig. 2a), the instructor prompted the class to ask questions of the proof and then to address, "What do you see that is the same? What do you see that makes them different?" through a think-pair-share mechanism, which was structured such that Partner A had to share a similarity and Partner B had to share a difference. A whole class discussion then ensued with a public record of these noticings (see Fig. 2b).

Regarding student opportunities to engage in AMPA, comparing and contrasting public records of proofs was quite successful in all cycles. Students noticed all of the distinctions we anticipated including beginning the proof with elements in the domain or co-domain group, naming of the elements, and warrants. These structural differences anticipated later discussions of what assumptions were needed, validating and modifying the proof approaches, and attending to differences in proof frameworks. However, we note that as in the small group discussion described in "Adjusting Partner-Exchange Structures to Increase Participation and Decrease Status Disparity During Proof Comprehension Phases" section, the proof approaches that became the focus of the class may have reinforced status differences. The type and focus of participation was catered towards two students who wrote and explained

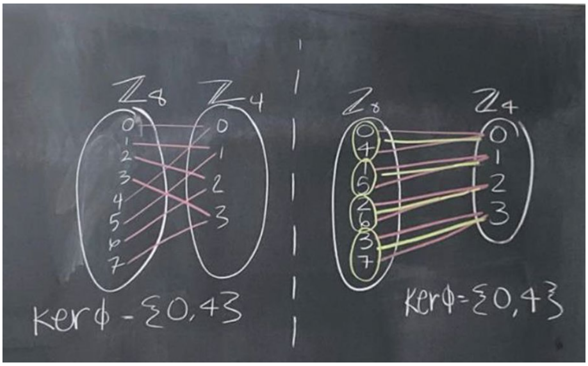

Fig. 3 Cycle 5 small group student boardwork illustrating the FIT

more complete deductive arguments. In later implementations, we changed the focus to the pre-existing student proofs to avoid this inadvertent difference in opportunity.

The attention to *whose* public record was also a factor in the other lessons. These records included not just proofs, but also examples to compare and generalize from. Although these examples were intended to be group products, we found they often reflected a particular individual's contributions. In order to address this issue, we engineered the group work structures to provide particular roles for each group member (Cycle 5 and 6). For example, when illustrating the FIT, students were given a set of chalk and each of the following roles was assigned to a different group member: identifying the domain and codomain elements, putting in the homomorphism map lines, identifying the kernel, and introducing lines representing the homomorphism. In this way, the public records (see Fig. 3) to be compared represented joint, rather than individual efforts.

Attention to the origins of public records was not attended to in the lab setting as the small number of students minimized the need for selecting only certain records to be shared. When transitioning to full class (cycle 3 and cycle 4), this issue became part of design minicycle discussions leading to the development of these two ways of countering the issue: introducing student work from outside of the class and introducing roles so all group members contribute to a public record.

Incorporating Public Records of Partial Information to Promote Access to Formal Ideas in Relation to Students' Informal Ideas

A major theme across all of HLTPS across implementations was navigating between formal and informal representations. One main use of public records of student thinking was to have a series of examples publicly available to make comparisons and support students in seeing important structural attributes to support rich discussions, develop understanding of theorems, and anticipate elements of proofs. In Cycle 5, the first in-person implementation of FIT, we found that the conversation that ensued in the whole class did not achieve the primary goal of generalizing across examples. The instructor guided a discussion focused on one example (shown in Fig. 3) that largely reflected a traditional Initiate-Response-Evaluate (IRE) pattern of discussion. That is, the instructor would ask a targeted question such as, "Where is the homomorphism?" and a student would respond "the red lines." The instructor then would endorse and elaborate. While the examples seemed to serve a productive role in small group time, we were not able to evidence engagement in AMPA during the larger discussion.

By contrast, in Cycle 6, the instructor made the decision to orchestrate this discussion to target a formal idea needed to make sense of the FIT proof. After students anticipated what needs to be proved for the theorem, the instructor tasked students with trying to write a definition for β by filling in: β (____) = ____. With a little discussion, students contributed that cosets go in the parentheses and they can be called "aK." However, identifying the corresponding output involved more work. One student suggested the "image of ϕ" (the relevant set) without specifying where a particular element goes. The instructor prompted students to compare across three examples on the board to try and use commonalities to identify a way to label the output of this isomorphism. After some initial student suggestions and discussion, they arrived at $\phi(a)$. The instructor then asked "Would that be consistent with what the three groups did?" walking around the room to consider each example. We highlight this implementation decision because it both supported students in AMPA (both in the moment and ultimately providing an access point in the proof comprehension activity) and helped the comparison of public records realize their full potential. We suggest an important transferable element: a partial formalization to connect between the informal examples and formal statement and proof. As discussed in earlier sections, students were usually able to work quite productively within an informal or formal system, but going between the two required some instructional intervention. We suggest recording a partial piece of information in formal symbolic form provides a means of scaffolding discussion such that a generalization could be anticipated in a way that would connect to the formal proof. Similar scaffolding occurred in other lessons such as the instructor providing partial diagrams for students to complete to make sense of the structural property proofs or engaging students in matching informal conjectures to formal statements of lemmas for Lagrange's Theorem. That is, providing partial information can support students in working between formal and informal systems while still providing opportunity for students to engage in AMPA.

Reflection on Selecting and Working with Public Records of Student Ideas in a Proof Based Setting

Overall, we found the use and comparison of public records to be a useful mechanism for focusing students on important structural elements that may have otherwise been hidden. In bringing this HLTP to the proof-based settings, we had to be especially cognizant of amplifying status issues. In a traditional lecture class, the instructor is the primary proof constructor and the proofs students engage with come from them. When creating a student-centered environment, there is a danger where a select few students generate the proofs which can reinforce status hierarchies. Further, the products of group work are not always reflective of all group members. To address these issues, we intentionally modified our group work instructions to better equalize the types of participation students engaged in and to increase the likelihood that public products contain elements from each group member. Our second major modification stemmed from the continued theme of navigating between informal and formal mathematics: the introduction of partially supplied formal information. With the set of examples illustrating the FIT, the initial implementation left these informal representations largely in isolation from the formal activity. In the second implementation, the instructor dedicated time to making one important informal to formal transition: defining the isomorphism. We also found this type of move important in the Lagrange Theorem task where students were positioned to notice generalities across their examples and arrive at key coset lemmas, described informally. The formalization process became a major transition between the examples and construction of the proof. Thus, we suggest that public records can serve two key roles: highlighting structural differences in proofs and promoting attention to structural generalities that can be formalized to anticipate continued AMPA in the more formal representation system.

Discussion

In this section, we revisit the set of HLTPs, provide an overview of how they differed from their K-12 counterparts, and then connect this work to the larger research base on inquiry and student-centered teaching in proof-based classrooms.

Expanding HLTPs from the K-12 Setting to the Proof Setting

Launching complex tasks served to provide common ground on mathematical terminology, promote access to opportunity to richly engage, and anticipate and emphasize important mathematical ideas and relationships. These fundamental components did not alter in the proof setting. However, in proof-based settings, the focus is on the theorem to be proven from a lens of meaning, logic, and anticipating proof structure. Further, the language and symbols have a high level of lexical complexity. Thus, there was a greater need to promote student attention to precision such as the role of quantifiers and making sense of the mathematical objects involved. Additionally, when the goal is to engage with the proof, it can be particularly challenging to identify key ideas (Raman, 2003) or use concepts and definitions (Moore, 1994). Complex task launch can support attention to important structures and provide the needed tools in terms of concepts and definitions to support proof construction, comprehension, and validation. Instructors and designers in proof-based classes may want to plan for when and how they will ask for greater precision around quantifiers and referents. This could involve planning tasks or questions where students are asked about quantities or are asked to explain what type of object is being referenced by certain symbols. Additionally, task launch serves not just to support access to mathematically dense symbols and language, but also anticipate structures ahead. While the most obvious way to engage students is by having them address formal definitions, we also found that supporting students in using key ideas towards proof required intentional discussion and linking between formal ideas and the informal structures needed. A parallel example might be found in analysis where intuition around limits often evokes attention to the independent variable first, whereas the formal definition needed for proving involves addressing the dependent variable first (see Swinyard & Larsen, 2012). We suggest instructors and designers explicitly consider how closely informal explorations reflect proof structures and find appropriate ways to bridge between student intuition and the needed structures to produce a proof.

Setting up and managing group work provided opportunity for students to engage with the mathematical task, encouraged equity of participation, and positioned students as contributors to mathematics. The commonalities across these instantiations were: a clear proof activity purpose (constructing, comprehending, validating) and the expectation, via instructor-provided structures, that all students participate and communicate about the mathematics. Ultimately, the iterations of this HLTPs increased in structure to better combat status issues that are often amplified in a setting where constructing a formal proof is valorized. If we compare to the K-12 setting, our structuring contained many similarities with complex instruction (e.g., Cohen, 1994); however, the structure and roles were often tied to the formal representation system. Of particular importance was supporting students traversing between the formal and informal to deconstruct existing proofs or to create examples that can be formalized. We suggest that instructors and designers consider not just the quality of a task created, but the ways they will engage students such that individual accountability and interdependence are required. The context of proofs can serve to exclude students who do not feel as confident with the abstract setting (Weber & Melhuish, 2022). Planning might include considering how students can engage productively in ways that are not just producing formal proofs, and how roles and responsibilities can be subdivided and assigned in relation to meaningful activity such as leading sense-making of parts of existing proofs or developing examples of ideas.

Selecting and working with public records of student ideas positioned students as contributors to the mathematical agenda, introduces resources into the common set of ideas, arguments, and representations for students to access, and engaged students in analyzing, critiquing and noticing important aspects of each other's mathematics. In the proof setting, we have focused this HLTP on comparing student proof approaches (which can lead to noticing differences in proof structures and arguments) and supporting students in generalizing and connecting to proofs. Specific to the proof setting was creating a task where students may viably create different proof frameworks (Selden & Selden, 1995). As in the K-12 setting, we found that comparison provided an opportunity for students to notice structural features that may be otherwise missed – the structure of a proof framework is one of the most important new elements in this setting. We also assigned groups different examples and intentionally focused on examples where their commonalities could be noticed and formalized to develop theorems and proofs. This involved unpacking what were either key ideas (Raman, 2003) or particular notation that would be found in later proofs (as in the Lagrange and FIT lessons). Students connected the records to support later formalization. Instructors and designers may consider what types of theorems have multiple approaches that may align themselves with meaningful structural comparisons. Additionally, they may plan for students to create examples whose comparison can support generalizations and connection to formal ideas. We also suggest consideration to whose ideas make it to a public space, how group products might be jointly created, and to include public records that are not solely formal proof productions.

Connecting to the Larger Literature Base in Proof-Based Courses

We have argued that the primary contribution of this paper is to draw explicit attention to HLTPs in a proof-based setting as well as to share lessons we learned about how to implement them effectively towards participation goals. We oriented this examination using our access and engagement heuristics. While we did not identify other design-based research articles (or empirical articles more generally) with such focus, we can make connections to other literature on curriculum and instruction in the advanced mathematics setting.

If we turn to the complex task launch, we can find practitioner reflections such as Reinholz (2020) who shared ways that their graduate analysis course launched tasks to include demonstrating the mechanics of the task, offering sets of questions, and providing instructions on possible next steps. We can also make connections to inquiry-oriented curricula that are driven by Realistic Mathematics Education. These tasks find their groundings in "experientially real" contexts where access is maximized (for example, see Larsen's (2013) trajectory for the guided reinvention of groups). There are also components of launching complex tasks in proof-contexts such as in Samkoff and Weber's (2015) proof comprehension strategies where students explore theorems and identify important definitions that may anticipate a proof. Additionally, pedagogical objects such as transformational records (Rasmussen & Marrongelle, 2006) can serve an essential role in supporting students in further mathematical activity. Our analysis adds some key insights about how to manage and adjust these kinds of tools to maintain the ambitious goals of engaging students in AMPA as well as ensuring all students have adequate resources to participate in the classroom activity.

We have found little literature about structuring and managing group work beyond the think-pair-share mechanism (see the *MAA Instructional Guide,* Abell et al., 2018) despite "managing group work" being one of the key roles of the instructor in inquiry classes (Ernst et al., 2017). More focus is placed on instructor discussion and intervention with small groups (e.g., Remillard, 2014) rather than embedded participation structures. Small group work serves an essential role in a number of the curriculum-based studies (e.g., Larsen et al., 2013) and intervention studies (e.g., Cilli-Turner, 2017). One mechanism that has been discussed is peer review (Reinholz & Pilgrim, 2021) or collaborative review (Cilli-Turner, 2017). These structures involve students constructing proofs and then sharing them with a partner or small group for critique and revision. In the case of Cilli-Turner, she expressed the challenge of having students direct comments and questions to each other rather than to the instructor. Others have endorsed norms such as Furinghetti et al. (2001) stating, "Doing collaborative group work means that students must be aware of the fact that everyone can and should contribute to the solution of the problem, and that sharing and comparing strategies and ideas is much more productive than working alone" (p. 232). However, they found that wanting this norm did not ensure all small groups worked collaboratively. Thus, we suggest that some of our structures for group work (developing expertise on a particular approach, leading discussion on specific proof comprehension questions) may be of use to researchers aiming for collaborative proof classrooms that meet this norm.

Working with public records of student proofs and thinking is another well-documented component of inquiry instruction. For example, IBL often relies on students presenting proofs and the class critiquing them (Starbird, 2015). Implementing inquiry-oriented curricula often involves intentional selection of student ideas to move along a desired progression (see Andrews-Larson et al., 2019; Lockwood et al., 2013). Curricular supports may include specific student ideas to look for and ideas to focus on during this discussion (Lockwood et al., 2013). Recent studies on instructors taking up these curricula have pinpointed initial shifts in how instructors' scaffold and select student ideas after teaching the class repeatedly (e.g., Andrews-Larson et al., 2019) noting a move away from primarily selecting correct responses for discussion. Further, Blanton and Stylianou (2014) have illustrated the instructor role in promoting students in reasoning about their own and others ideas once ideas are available for discussion. We further this work with considerations of *whose* ideas are made public and ways that student work might be leveraged in a *formal proof setting* such as to compare proof approaches or generalize and formalize key conjectures.

Conclusions, Limitations, and Future Research

The biggest lessons we learned from our design project are (1) challenges involved in moving between formal and informal representational systems impacted nearly every HLTP implementation and (2) equitable participation did not occur by attending only to access and opportunity. Overall, the adaptation of the HLTPs to the proof-based setting were not substantially different than their K-12 counterparts. However, the formal and abstract setting necessitated precision around objects, language, and quantification. There was a need for instructor guidance and task features to support traversal between the formal and non-formal representation systems throughout implementation of all the HLTPs. Otherwise, students often engage in the formal representation system without drawing upon their informal explorations with examples or diagrams. In terms of participation, we found structuring group work to distribute responsibility and providing opportunity beyond just formal proof production as essential. As design researchers make shifts from interview setting to classrooms, we suggest they consider (1) ways that their tasks can be partnered with initial structuring to allow for more equitable participation and (2) ways to mitigate a few students' public records being centered via expansion in their creation or types.

Finally, we note the HLTPs cannot be disentangled from the context and nature of the tasks in this project. Thus, while we contribute instructional elements that we see as generalizable, many may not be usable in courses that do not share the common proof construction, validation, and construction objectives. Future research may consider how other tasks and subject areas may elicit different elements of the HLTPs in proof-based settings. Additionally, while we attended broadly to access and engagement in activity, we did not take a critical lens to our analyses. Further research could consider how student identity may be reflected in whose activity is elicited and valorized.

Appendix

	Lesson 1: Structural Property Task	Lesson 2: Lagrange's Theorem Task	Lesson 3: First Isomorphism Theorem Task
Primary Proof Activity	Proof Validating	Proof Constructing	Proof Comprehending
Proof Learning	Role of conclusion in proof framework	Using a diagram exploration to identify a key idea	Attending to global and local aspects when reading a proof
Outline of Lesson Structure	**Opening Discussion** • Public record of assumptions and conclusion • Discuss key definitions **Discussion of Proof Approaches in Small Groups** • Develop expertise around one of two student proofs • Explain proof to partner who is prompted to state one thing that makes sense and one question **Public Discussion of Two Approaches** • Presentation • Identifying similarities and differences (think-pair-share) • Public Record of similarities and differences **Proof and Statement Analysis** • Conjecture what assumptions are needed based on the existing proofs • Public Record of conjectured statements (with varying properties: 1–1 and onto) • Testing statements use proofs and examples to determine what properties are needed **Counterexample to Identify the Necessity of Onto (Visual Representation)** • Function diagram discussion showing the role of onto **Summary and Conclusion** • Finalizing of revised statement (onto, but not 1–1) • Discussion of patching the proof that did not use onto • Discussion of the role of conclusion in structuring proofs (proof framework)	**Opening Discussion and Exploration** • Exploring examples groups and the order of their subgroups to generate a conjecture (in small groups) • Creating a public record of conjectures from different groups • Connecting conjecture to formal Lagrange Statement • Formally defining divisibility and exploring the meaning of "multiplication" on boards and coming to class consensus (to anticipate proof structure) **Creating Cosets in Small Groups** • Each group works with a different group (and subgroup) to create cosets in a form that can be reasoned with diagrammatically **Conjecture Discussion** • Discussion of noticings and conjectures about the structure of the cosets to arrive at key lemmas for the proof of Lagrange's Theorem **Matching Class Lemmas to their Formalizations in Small Groups** • A set of six formal statements to identify as the translation of an informal lemma or a tool to prove one of the lemmas **Proving Lagrange's Theorem** • Small group and whole class discussion of structuring Lagrange's Proof using the lemmas **Summary and Conclusion** • Discussing the role of the key idea from the coset diagram examples • Wrap-up on the implications of Lagrange's Theorem	**Opening Discussion** • Discussion of key concepts and definitions in the theorem **Small Group Exploration of Specific Examples** • Each small group works at board space to connect the theorem to a specific example using a function diagram (and assigned roles) **Class Discussion of Defining the Isomorphism Map** • Identifying a map in symbolic form that will describe the input (cosets) and outputs (image of the coset representative) that is consistent across the examples **Discussion (small group and whole class) of Proof Structure** • Identifying what needs to be proven • Subdividing the proof to find the sections of what needs to be proven **Making sense of a subsection of a proof** • Each small group is responsible for one of four sections. Each member of a group has a question for leading discussion **Class Presentations** • A representative from each small group explains their section to the class **Summary and Conclusion** • Summarizing the proof at a high level • Discussing the practice of proof comprehension

References

Abell, M., Braddy, L., Ensley, D., Ludwig, L., & Soto-Johnson, H. (2018). *MAA Instructional Practices Guide*. Mathematical Association of America.

Andrews-Larson, C., McCrackin, S., & Kasper, V. (2019). The next time around: Scaffolding and shifts in argumentation in initial and subsequent implementations of inquiry-oriented instructional materials. *The Journal of Mathematical Behavior, 56*, 100719.

Andrews-Larson, C., Wawro, M., & Zandieh, M. (2017). A hypothetical learning trajectory for conceptualizing matrices as linear transformations. *International Journal of Mathematical Education in Science and Technology, 48*(6), 809–829.

Ball, D. L., Sleep, L., Boerst, T. A., & Bass, H. (2009). Combining the development of practice and the practice of development in teacher education. *The Elementary School Journal, 109*(5), 458–474.

Blanton, M. L., & Stylianou, D. A. (2014). Understanding the role of transactive reasoning in classroom discourse as students learn to construct proofs. *The Journal of Mathematical Behavior, 34*, 76–98.

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record, 110*(3), 608–645.

Bouhjar, K., Andrews-Larson, C., & Haider, M. Q. (2021). An analytical comparison of students' reasoning in the context of Inquiry-Oriented Instruction: The case of span and linear independence. *The Journal of Mathematical Behavior, 64*, 100908.

Brown, S. (2018). E-IBL, proof scripts, and identities: An exploration of theoretical relationships. In A. Weinberg, C. Rasmussen, J. Rabin, M. Wawro, & S. Brown (Eds.), Proceedings of the 21st Annual Conference on Research in Undergraduate Mathematics Education (pp. 1–15). Mathematical Association of America, SIGMAA on RUME.

Brown, R. (2009). Teaching for social justice: Exploring the development of student agency through participation in the literacy practices of a mathematics classroom. *Journal of Mathematics Teacher Education, 12*(3), 171–185.

Cilli-Turner, E. (2017). Impacts of inquiry pedagogy on undergraduate students conceptions of the function of proof. *The Journal of Mathematical Behavior, 48*, 14–21.

Cobb, P., & Gravemeijer, K. (2014). Experimenting to support and understand learning processes. In *Handbook of design research methods in education* (pp. 86–113). Routledge.

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64*(1), 1–35.

Cohen, E. G., & Lotan, R. A. (1997). *Working for equity in heterogeneous classrooms: Sociological theory in practice*. Teachers College Press.

Dawkins, P. C., Oehrtman, M., & Mahavier, W. T. (2019). Professor goals and student experiences in traditional IBL real analysis: A case study. *International Journal of Research in Undergraduate Mathematics Education, 5*(3), 315–336.

Dawkins, P. C., & Weber, K. (2017). Values and norms of proof for mathematicians and students. *Educational Studies in Mathematics, 95*(2), 123–142.

Durkin, K., Star, J. R., & Rittle-Johnson, B. (2017). Using comparison of multiple strategies in the mathematics classroom: Lessons learned and next steps. *ZDM Mathematics Education, 49*(4), 585–597.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction, 20*(4), 399–483.

Ernst, D. C., Hodge, A., & Yoshinobu, S. (2017). What is inquiry-based learning. *Notices of the AMS, 64*(6), 570–574.

Esmonde, I. (2009). Ideas and identities: Supporting equity in cooperative mathematics learning. *Review of Educational Research, 79*(2), 1008–1043.

Fukawa-Connelly, T. P. (2012). A case study of one instructor's lecture-based teaching of proof in abstract algebra: Making sense of her pedagogical moves. *Educational Studies in Mathematics, 81*(3), 325–345.

Furinghetti, F., Olivero, F., IV., & Paola, D. (2001). Students approaching proof through conjectures: Snapshots in a classroom. *International Journal of Mathematical Education in Science and Technology, 32*(3), 319–335.

Garuti, R., Boero, P., & Lemut, E. (1998). Cognitive Unity of Theorems and Difficulty of Proof. *Proceedings of the international group for the psychology of mathematics education PME-XXII*, vol. 2, (pp. 345–352). Stellenbosch.

Gillies, R. M. (2003). Structuring cooperative group work in classrooms. *International Journal of Educational Research, 39*(1–2), 35–49.

Hazzan, O. (1999). Reducing abstraction level when learning abstract algebra concepts. *Educational Studies in Mathematics, 40*(1), 71–90.

Herbel-Eisenmann, B. A. (2002). Using student contributions and multiple representations to develop mathematical language. *Mathematics Teaching in the Middle School, 8*(2), 100.

Hicks, M. D., Tucci, A. A., Koehne, C. R., Melhuish, K. M., & Bishop, J. L. (2021). Examining the Distribution of Authority in an Inquiry-Oriented Abstract Algebra Environment. In S. S. Karunakaran, & A. Higgins (Eds.), *2021 Research in Undergraduate Mathematics Education Reports*.

Hlas, A. C., & Hlas, C. S. (2012). A review of high-leverage teaching practices: Making connections between mathematics and foreign languages. *Foreign Language Annals, 45*(s1), s76–s97.

Jackson, K., & Cobb, P. (2010). Refining a vision of ambitious mathematics instruction to address issues of equity. Paper presented at the annual meeting of the American Educational Research Association, Denver, CO.

Jackson, K., Garrison, A., Wilson, J., Gibbons, L., & Shahan, E. (2013). Exploring relationships between setting up complex tasks and opportunities to learn in concluding whole-class discussions in middle-grades mathematics instruction. *Journal for Research in Mathematics Education, 44*(4), 646–682.

Jackson, K. J., Shahan, E. C., Gibbons, L. K., & Cobb, P. A. (2012). Launching complex tasks. *Mathematics Teaching in the Middle School, 18*(1), 24–29.

Johnson, E., Andrews-Larson, C., Keene, K., Melhuish, K., Keller, R., & Fortune, N. (2020). Inquiry and gender inequity in the undergraduate mathematics classroom. *Journal for Research in Mathematics Education, 51*(4), 504–516.

Kaddoura, M. (2013). Think pair share: A teaching learning strategy to enhance students' critical thinking. *Educational Research Quarterly, 36*(4), 3–24.

Khisty, L. L., & Chval, K. B. (2002). Pedagogic discourse and equity in mathematics: When teachers' talk matters. *Mathematics Education Research Journal, 14*(3), 154–168.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*(1), 29–63.

Larsen, S., Johnson, E., & Weber, K. (Eds.) (2013). The teaching abstract algebra for understanding project: designing and scaling up a curriculum innovation. *Journal of Mathematical Behavior, 32*(4).

Larsen, S. P. (2013). A local instructional theory for the guided reinvention of the group and isomorphism concepts. *The Journal of Mathematical Behavior, 32*(4), 712–725.

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

Laursen, S. L., & Rasmussen, C. (2019). I on the prize: Inquiry approaches in undergraduate mathematics. *International Journal of Research in Undergraduate Mathematics Education, 5*(1), 129–146.

Leron, U., & Zaslavsky, O. (2013). Generic proving: Reflections on scope and method. *For the Learning of Mathematics, 33*(3), 24–30.

Lew, K., & Mejía-Ramos, J. P. (2019). Linguistic conventions of mathematical proof writing at the undergraduate level: Mathematicians' and students' perspectives. *Journal for Research in Mathematics Education, 50*(2), 121–155.

Livers, S. D., & Bay-Williams, J. M. (2014). Vocabulary support: Constructing (not obstructing) meaning. *Mathematics Teaching in the Middle School, 20*(3), 152–159.

Lockwood, E., Johnson, E., & Larsen, S. (2013). Developing instructor support materials for an inquiry-oriented curriculum. *The Journal of Mathematical Behavior, 32*(4), 776–790.

Lotan, R. A. (2003). Group-worthy tasks. *Educational Leadership, 60*(6), 72–75.

McClain, K., & Cobb, P. (1998). The role of imagery and discourse in supporting students' mathematical development. In M. Lampert & M. Blunk (Eds.), *Talking Mathematics in School: Studies of Teaching and Learning* (pp. 56–81). Cambridge University Press.

Mejía-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.

Mejía-Ramos, J. P., & Inglis, M. (2009). What are the argumentative activities associated with proof? *Research in Mathematics Education, 11*(1), 77–78.
Melhuish, K., Thanheiser, E., Heaton, R., Sorto, A. Strickland, S., & Sugimoto, A. (2020). The Math Habits Tool - Research Version [Measurement instrument] Retrieved 2020, from http://mathhabits.wp.txstate.edu
Melhuish, K., Fukawa-Connelly, T., Dawkins, P. C., Woods, C., & Weber, K. (2022a). Collegiate mathematics teaching in proof-based courses: What we now know and what we have yet to learn. *The Journal of Mathematical Behavior, 67*, 100986.
Melhuish, K., Larsen, S., & Cook, S. (2019). When students prove a theorem without explicitly using a necessary condition: Digging into a subtle problem from practice. *International Journal of Research in Undergraduate Mathematics Education, 5*(2), 205–227.
Melhuish, K., Vroom, K., Lew, K., & Ellis, B. (2022b). *Operationalizing Authentic Mathematical Proof Activity Using Disciplinary Tools [Manuscript submitted for publication]*. Texas State University.
Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics, 27*(3), 249–266.
Moschkovich, J. (2013). Principles and guidelines for equitable mathematics teaching practices and materials for English language learners. *Journal of Urban Mathematics Education, 6*(1), 45–57.
Nardi, E. (2000). Mathematics undergraduates' responses to semantic abbreviations, 'geometric' images and multi-level abstractions in group theory. *Educational Studies in Mathematics, 43*(2), 169–189.
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.
Rasmussen, C., & Kwon, O. N. (2007). An inquiry-oriented approach to undergraduate mathematics. *The Journal of Mathematical Behavior, 26*(3), 189–194.
Rasmussen, C., & Marrongelle, K. (2006). Pedagogical content tools: Integrating student reasoning and mathematics in instruction. *Journal for Research in Mathematics Education, 37*(5), 388–420.
Rasmussen, C., Wawro, M., & Zandieh, M. (2015). Examining individual and collective level mathematical progress. *Educational Studies in Mathematics, 88*(2), 259–281.
Reinholz, D., Johnson, E., Andrews-Larson, C., Stone-Johnstone, A., Smith, J., Mullins, S. B., Fortune, N., Keene, K., & Shah, N. (2022). When active learning is inequitable: Women's participation predicts gender inequities in mathematical performance. *Journal for Research in Mathematics Education*.
Reinholz, D. L. (2020). Five practices for supporting inquiry in analysis. *Primus, 30*(1), 19–35.
Reinholz, D. L., & Pilgrim, M. E. (2021). Student sensemaking of proofs at various distances: The role of epistemic, rhetorical, and ontological distance in the peer review process. *Educational Studies in Mathematics, 106*(2), 211–229.
Remillard, K. S. (2014). Identifying discursive entry points in paired-novice discourse as a first step in penetrating the paradox of learning mathematical proof. *The Journal of Mathematical Behavior, 34*, 99–113.
Samkoff, A., & Weber, K. (2015). Lessons learned from an instructional intervention on proof comprehension. *The Journal of Mathematical Behavior, 39*, 28–50.
Saxe, K., & Braddy, L. (2015). *A Common vision for undergraduate mathematical sciences programs in 2025*. Mathematical Association of America.
Selden, A., & Selden, J. (2017). A comparison of proof comprehension, proof construction, proof validation and proof evaluation. In *Proceedings of the Conference on Didactics of Mathematics in Higher Education as a Scientific Discipline* (pp. 339–345).
Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics, 29*(2), 123–151.
Silver, E. A., & Stein, M. K. (1996). The QUASAR project: The" revolution of the possible" in mathematics instructional reform in urban middle schools. *Urban Education, 30*(4), 476–521.
Simpson, A. (2015). The anatomy of a mathematical proof: Implications for analyses with Toulmin's scheme. *Educational Studies in Mathematics, 90*(1), 1–17.
Smith, M. S., Bill, V., & Hughes, E. K. (2008). Thinking through a lesson: Successfully implementing high-level tasks. *Mathematics Teaching in the Middle School, 14*(3), 132–138.
Spooner, F., Saunders, A., Root, J., & Brosh, C. (2017). Promoting access to common core mathematics for students with severe disabilities through mathematical problem solving. *Research and Practice for Persons with Severe Disabilities, 42*(3), 171–186.

Staples, M. (2007). Supporting whole-class collaborative inquiry in a secondary mathematics classroom. *Cognition and Instruction, 25*(2–3), 161–217.
Starbird, M. (2015). Inquiry-based learning through the life of the MAA. *A Century of Advancing Mathematics, 81*, 239.
Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*(4), 313–340.
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* (pp. 237–266). National Council of Teachers of Mathematics.
Sullivan, P., Zevenbergen, R., & Mousley, J. (2003). The Contexts of mathematics tasks and the context of the classroom: Are we including all students? *Mathematics Education Research Journal, 15*(2), 107–121.
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.
TeachingWorks. (2018). High leverage practices. http://www.Teachingworks.org
Thanheiser, E., & Melhuish, K. (2022). Teaching routines and student-centered mathematics instruction: The essential role of: Conferring to understand student thinking and reasoning. [Manuscript submitted for publication]. Fariborz Maseeh Department of Mathematics + Statistics, Portland State University.
Toulmin, S. E. (1958). *The Uses of argument*. Cambridge University Press.
Wawro, M., Rasmussen, C., Zandieh, M., Sweeney, G. F., & Larson, C. (2012). An inquiry-oriented approach to span and linear independence: The case of the magic carpet ride sequence. *Primus, 22*(8), 577–599.
Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology, 79*(1), 1–28.
Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics, 48*(1), 101–119.
Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education, 39*(4), 431–459.
Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics, 56*(2), 209–234.
Weber, K., & Alcock, L. (2005). Using warranted implications to understand and validate proofs. *For the Learning of Mathematics, 25*(1), 34–51.
Weber, K., & Mejía-Ramos, J. P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics, 76*(3), 329–344.
Weber, K., & Melhuish, K. (2022). Can we engage students in authentic mathematical activity while embracing critical pedagogy? A commentary on the tensions between disciplinary activity and critical education. *Canadian Journal of Science, Mathematics and Technology Education, 22*(2), 305–314.
Wilburne, J., Polly, D., Franz, D., & Wagstaff, D. A. (2018). Mathematics teachers' implementation of high-leverage teaching practices: AQ-sort study. *School Science and Mathematics, 118*(6), 232–243.
Woods, D. M., & Wilhelm, A. G. (2020). Learning to launch complex tasks: How instructional visions influence the exploration of the practice. *Mathematics Teacher Educator, 8*(3), 105–119.
Zazkis, D., Weber, K., & Mejía-Ramos, J. P. (2016). Bridging the gap between graphical arguments and verbal-symbolic proofs in a real analysis context. *Educational Studies in Mathematics, 93*(2), 155–173.