Rethinking doctoral qualifying exams and candidacy in the physical sciences: Learning toward scientific legitimacy

Román Liera,[1] Aireale J. Rodgers,[2] Lauren N. Irwin,[3] and Julie R. Posselt[4]

[1]*Department of Educational Leadership, Montclair State University, Montclair, New Jersey 07043, USA*
[2]*Center for the Humanities, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA*
[3]*Educational Leadership & Policy Studies, University of Tennessee, Knoxville, Tennessee 37996, USA*
[4]*Rossier School of Education, University of Southern California, Los Angeles, California 90089, USA*

(Received 23 March 2023; accepted 11 July 2023; published 14 August 2023)

There is growing awareness that established structures of higher education are often predicated on problematic assumptions about merit, excellence, and rigor. Doctoral qualifying exams, for example, are required to advance to candidacy in many Ph.D. programs despite decades of documented concerns about the implications of standard modes for student equity and well-being. As more Ph.D. programs move to reform these exams and candidacy requirements, it is important to understand how Ph.D. programs, as academic organizations, construct the significance of the qualifying exam. A sociocultural lens suggests qualifying exams and the learning that enables their passage are symbolic rituals that move doctoral students from legitimate peripheral participation toward full membership and belonging in academic communities of practice. We conducted a comparative case study to understand how two Ph.D. programs in the physical sciences that have reformed their candidacy requirements—one elite and one middle ranked but striving for respect—constructed the significance and purpose of their qualifying exam and the broader transition to candidacy. Our inquiry included the contexts and mechanisms that mediated student learning. Through interviews with faculty, staff, and students, we found that the Ph.D. programs' recognition of their status within their respective disciplines emerged as a crucial component in constructions about the significance of exams and candidacy. The middle-ranked Ph.D. program changed the exam and candidacy structure to reflect legitimate practices in their discipline. The elite Ph.D. program created multiple pathways toward candidacy to mitigate long-standing concerns about gender equity and student well-being. Despite the structural changes, the Ph.D. programs left intact cultural understandings of merit, excellence, and rigor that maintain inequity in doctoral socialization. Our findings suggest that researchers and practitioners should pay more attention to designing and implementing structures that facilitate faculty assessments of doctoral student learning.

DOI: 10.1103/PhysRevPhysEducRes.19.020110

I. INTRODUCTION

Among gatekeeping processes in higher education, doctoral qualifying exams[1] and other candidacy requirements are understudied structures that have implications for degree completion, equity, and student well-being. Conventional wisdom suggests that faculty use these exams to assess knowledge and skills necessary for independent dissertation research [1–3]. Yet, a growing number of faculty, staff, and students in the physical sciences recognize that traditional qualifying exams deviate from the work expected of scholars, are a significant point of student attrition, and may threaten student well-being [4–6]. A comprehensive review of published literature finds the "need for more evidence-based research on qualifying exams, which are nearly ubiquitous as a major doctoral training milestone," in order to generate "evidence for best practices" [7] (p. 19).

There have also been field-specific efforts to adopt evidence-based, learning-focused approaches to qualifying exams in chemistry [8] and physics [4]. Recent studies have critiqued typical exams in these fields for their low passing rates [9], unrealistic time constraints, unclear expectations and procedures [4], and apparent apathy among faculty to mentor and prepare students [10]. In some instances, Ph.D. programs in physics have modified and changed the qualifying exam process to improve the student experience [4]. In both chemistry and physics, attention to qualifying

[1]Qualifying exams, comprehensive exams, and preliminary exams are variously used in different programs and disciplines; our use of qualifying exams is specifically in reference to exams that are required as a condition of advancement to candidacy.

exams is part of broader, national conversations about pressing cultural issues in graduate education [4,8].

To inform these discussions and associated reform efforts, we report a comparative case study of physical science Ph.D. programs that have thoughtfully redesigned their exams and other candidacy requirements. We used a sociocultural theoretical framework that is attuned to learning, legitimacy, and status dynamics [11] to understand how programs may not only (re)design requirements with an eye to "preparation for independent scholarship" but, one level deeper, with attention to the qualities of scientists whom they hope to develop and attention to program status and legitimacy. We sought to answer two research questions: (1) How do students, faculty, and Ph.D. programs in the physical sciences construct the significance and purposes of doctoral qualifying exams? (2) What contexts and mechanisms mediate learning for the exam and the larger candidacy transition?

II. LITERATURE REVIEW

The literature on qualifying exams has two major streams: their purposes and pitfalls for learning and subjective assessment in qualifying exams. We review this literature as a foundation for our research, including providing critiques and new directions.

A. The purposes and pitfalls of qualifying exams

Faculty instituted qualifying exams in the early 20th century out of concern that the dissertation defense "had decayed into an exercise serving only a ritualistic function" [12] (p. 399). Intending to be an "assurance of quality" for the Ph.D. [13] (p. 79), faculty designed qualifying examinations to assess subject matter knowledge including topics, theories, and methods [2,14]. For example, a report by the joint AAPT-APS Task Force on Graduate Education in Physics [15] found that 90% of physics departments that required qualifying exams covered quantum mechanics and classical electrodynamics. Posselt and Liera [16] collected survey data to explore variations in structures and perceived purposes of doctoral candidacy exams in chemistry ($n = 40$ Ph.D. programs) and physics ($n = 42$ Ph.D. programs). They found that 85% of respondents in chemistry and 79% in physics perceived the primary purpose of qualifying exams to be preparing students for future activities. In both fields, 50%–54% of programs named ensuring a rigorous Ph.D. program as the second most important purpose of qualifying exams.

A standardized evaluation of each student's breadth of disciplinary knowledge has been assumed to offer programs two benefits. First, simply requiring them would aid programs in making a case for the rigor of their training [17]. Additionally, student failure on the exam would give programs a legitimized rationale to dismiss "weaker" students before making a substantive investment in their dissertation research [12]. In short, qualifying exams confer legitimacy on programs that require them and the students who pass them [17].

For students, the literature suggests qualifying exams may also have developmental dimensions. The preparation process "allows graduate students to feel accomplished and knowledgeable in their discipline," according to one study [5] (p. 26). Other developmental outcomes include mastering a body of knowledge and learning research design and technical skills necessary for the dissertation [1,14,18,19]. In chemistry, Harshman argued that faculty implemented the qualifying exam for students to demonstrate not only such disciplinary knowledge but also their independent thinking [3].

Yet, insufficient support for these developmental processes and preparing students for the exams can threaten well-being, socializing doctoral students to view stress and anxiety as inherent in scholarly life and achievement [5]. Qualifying exams have been described in the literature as a rite of passage [12], "an obstacle course and ritual gauntlet" [20] (p. 30), a "hurdle" to be surmounted [21] (para. 3), and an ambiguous terrain where students can lose their way [22]. A systematic grounded theory from 125 interviews with doctoral students about their experiences with the qualifying exam found that "factors causing stress differed from student to student," and across phases of the exam and transition to candidacy [22] (p. 68). In the preparation phase, ambiguous expectations of how much studying was enough and the need to drop other priorities created stress. The perceived stress of taking the exam depended on its structure,[2] which varied across programs and universities. Burnout often followed, tempered with relief, depending upon the outcome [22].

B. Subjectivity and opacity in qualifying exams

Unhealthy structures in doctoral education may persist because they fulfill important latent, cultural functions or because members have not questioned their underlying assumptions [23,24]. Our review identified a series of papers that describe unquestioned subjectivity and opacity as problematic foundations of most qualifying exams.[3] First, subjectivity is inherent in their design, including arbitrary judgments about acceptable exam formats and the subject matter viewed as worthy of assessing [4]. Many qualifying exams do not elicit the

[2]Posselt and Liera [16] found that chemistry and physics Ph.D. programs use a combination of comprehensive, qualifying, and preliminary exams for advancement to candidacy. Although the timing of these different exams varied across and within fields, it was not uncommon in physics to require a preliminary exam before a dissertation proposal.

[3]A tendency toward subjective judgment is consistent with research on the judgments that guide other academic gatekeeping contexts, including graduate admissions [23], faculty hiring [25,26], and peer review [27].

range of professional skills that students need outside of Ph.D. coursework [3,28].

A related concern surrounds a lack of specificity and clarity in what qualifying and other exams in the physical sciences evaluate [4]. Here, it appears that different assessment structures pose different problems. On written exams, nuances in what acceptable answers consist of may not be communicated to students. Outside the qualifying exam context, Marshman *et al.* found that even within the high consensus field of physics, instructors assess responses in quantum mechanics for conceptual reasoning, whereas they assess introductory physics solutions for "evidence of understanding" [29] (p. 10). Oral defenses "discriminate between superficial and real knowledge, [and] test examinees' abilities to synthesize and consolidate rather than merely regurgitate information" [18] (p. 177). In practice, professors may be attuned in defenses to competencies as widely varying as thinking on one's feet, convincingly communicating one's research, and/or contributing to the field [19,30]. Assessments of these behaviors may be especially subject to implicit biases [31]. Across formats, opacity also poses a learning problem in that students are shielded from information about what is understood to evince their learning.

C. Critique of the literature and new directions

In summary, scholars have documented the purposes of qualifying examinations, as well as issues that stem from ambiguous expectations, failure to prepare students systematically, and high stakes. Scholars have also identified problems with the foundations of exams: subjectivity and opacity in evaluation design and criteria, to name two. These problems are not new. As early as 1932, Pressey and colleagues [32] described the candidacy examination as the "final ordeal" because students were subjected to unclear standards and unreliable methods to scrutinize their behaviors in a setting that determined their educational and professional fate.

We also observed stagnation in the current literature, which sociocultural theoretical perspectives may help address. Most studies reinforce a surface-level, functionalist narrative that qualifying exams and the broader transition to candidacy use to assess for readiness for "independent scholarship" (e.g., [1]). In some places, sheer content knowledge demonstrated via problem solving may suffice for candidacy because faculty assume the scientific development of doctoral students happens in the lab. But in other places, as evidence on the range of behaviors evaluated in oral defenses suggests, cultural judgments are part of determining qualifying exams passage or granting candidacy. What counts as "readiness" may include alignment with skills and behaviors that legitimate scientists are assumed to have—or should be well on the path to developing. Sociocultural perspectives may attune us to performance on the qualifying exams as a matter of performing valued disciplinary cultural knowledge. Candidacy may also be culturally significant as a new membership status in one's department or discipline. Moreover, because sociocultural theory recognizes learning as situated and unfolding in specific contexts and communities, research from this angle can explore how learning and development may be structured and experienced in Ph.D. programs with different visions and designs for candidacy transition.

Organizational analysis can also bring fresh perspective. Scholars have acknowledged how the origins of qualifying exams in Ph.D. programs demonstrate rigor [17], but they have not investigated how programs as organizations design and implement requirements for students. This is a notable gap because such factors may also affect reform considerations. Further, to our knowledge, no studies have investigated satisfaction with reform or elimination of qualifying exams. We need to uplift these stories in the physical science community to highlight alternatives and challenge status-quo assumptions about what comprises legitimate training. There are more avenues for possible research than can be addressed in a single paper. Nevertheless, a sociocultural perspective on organizational efforts to reimagine the transition to candidacy can shed fresh light on the nature of learning in doctoral education as well as reforming or eliminating exams.

III. THEORETICAL FRAMEWORK: LEARNING LEGITIMACY

Sociocultural theory is marked by attention to the cultural contexts of social life. It looks beneath the surface of observable social phenomena to understand the meanings and assumptions that motivate people and their behaviors in a given context. It is therefore useful for examining learning for candidacy as "situated learning"—that is, learning that occurs in community cultural contexts marked by distinct and evolving programmatic structures and concerns, as well as wider and enduring discipline-level assumptions about what knowledge and skills are central and marginal [33,34]. Faculty and students alike have described the disciplinary culture in chemistry and physics Ph.D. programs as competitive [10,35,36]. We weave four core strands from sociocultural theory to study qualifying exams, the transition to candidacy, and their reimagination: legitimacy, socialization, communities of practice, and status.

A. Legitimacy

A sociocultural perspective suggests that one reason "independent scholarship" may persist as an ideal is that doctoral education is widely understood to be a matter of producing "stewards of the disciplines" [37]. Professors need to feel that they can trust the next generation to conduct research and carry themselves in ways that are aligned with recognized knowledge, practices, and norms. Congruence with community culture is a question of legitimacy [38,39], defined as alignment with established cultural beliefs and norms. Cultural alignment is especially important in status-oriented organizational sectors like education because it helps buy membership and belonging for individuals and organizations [40]. Given the socially constructed nature of legitimacy and its power to shape membership and opportunities, how emerging scholars learn what counts as legitimate warrants attention—as do the actions of organizations to construct processes for this learning.

B. Socialization

Socialization refers to the "processes through which [students] gain the knowledge, skills, and values necessary for successful entry into a professional career requiring an advanced level of specialized knowledge" [41] (p. iii). What gatekeepers deem "necessary for successful entry" is grounded in context specific—not objective or singular—beliefs about what scholars should be able to know and do, even within fields like physics that are relatively paradigmatic [42]. Differences in beliefs on these matters, along with the rarity with which educational organizations revisit inherited practices, may help explain the variety of exam structures and requirements that are currently in use [12,19]. Socialization to the knowledge and skills that are viewed as legitimate in one's discipline happens in community. As students engage over time in disciplinary practice, they become part of, and potentially central to, wider communities of practice.

C. Communities of practice

Communities of practice are groups who work toward common goals through shared interests and motivations [33]. This strand of sociocultural theory highlights how people are socialized to perform a community's legitimized culture in ways that move them from being outsiders to being insiders of the group. This transition includes a phase of "legitimate peripheral participation" that involves learning what the community values [33,43]. This learning equates with socialization. Students engage in joint activities and discussions to co-create knowledge, practices, and norms that the group values, adopts, and implements [43]. In physics, labs and courseworks are important sites for engagement with communities of practice, where interactions "create opportunities for vicarious learning" [44] (p. 1). In these settings, STEM students are exposed to a range (often limited) of potential trajectories for disciplinary identity and scholarly development [45,46]. Recognizing that the transition to doctoral candidacy marks a formal membership status, this theory is useful in describing (1) how communal activities associated with the qualifying exam and transition to candidacy may mediate (i.e., the environments and tools the learner has access to facilitates their learning and development) learning of legitimated knowledge, and (2) the cultural dynamics of learning that facilitate or impede disciplinary membership and belonging.

D. Status

We anticipated that organizational status might shape how Ph.D. programs and their members construct the purpose—and thus the necessary design—of candidacy requirements, as well as potential options for changing their typical approach. Just as individuals are part of status hierarchies, academic departments in the physical sciences are associated with national and international disciplinary communities, and some have more status than others. Rankings have been for decades a critical status indicator, although other factors affecting organizational status in disciplines include the production of faculty, positions in more or less elite colleges and universities, and signature intellectual contributions. Phillips and Zuckerman found a U-shaped pattern in organizational willingness to innovate based on their status [47]. High-status organizations were confident in deviating wildly from conventional expectations, and low-status organizations were also open to major changes because they had little to lose. However, middle-status organizations tended to be more cautious, choosing to innovate only on terms that they believed would be acceptable at the field level [47]. With this in mind, we anticipated that faculty at high-status and low-status Ph.D. programs would be more confident making dramatic changes to candidacy, while faculty at middle-status programs would redesign with an eye to their discipline's legitimized skills and knowledge.

In summary, concepts from sociocultural theory can be used to reinterpret how we read the espoused purposes of qualifying exams and transition to candidacy. Through this lens, readiness is a question of legitimacy; as one is socialized to a community of practice, one is more easily recognized as legitimate by other community members. The transition to candidacy can be thought of as a cultural ritual in which preparation and positive evaluation move doctoral students toward full membership in disciplinary and departmental communities of practice. Activities that surround the transition to candidacy are anticipated to be sites of learning formal subject matter knowledge and informal cultural behaviors (i.e., practice) that disciplinary members expect of one another [33]. Given a program's status in the discipline, candidacy requirements may vary, and reform may follow broadly predictable patterns.

IV. METHODOLOGY

Our research questions were as follows: (1) How do students, faculty, and Ph.D. programs in the physical

TABLE I. Characteristics of data collected by program and role.

	Western (chemistry)			Midwestern (physics)		
Data collection method	Interviews	Focus group	Total	Interviews	Focus group	Total
Faculty	6	⋯	6	4	⋯	4
Staff	1	⋯	1	1	⋯	1

sciences construct the significance and purposes of doctoral qualifying exams? and (2) What contexts and mechanisms mediate learning for the exam and the larger candidacy transition? Comparative case study facilitates analysis of patterns within and across bounded systems because it recognizes the importance of context in shaping phenomena [48,49]. This methodology was appropriate for examining and comparing how physical science Ph.D. programs (re)designed the transition to candidacy and structures for student learning. Moreover, the systems of Ph.D. programs for advancing students to candidacy were clear examples of bounded systems [48,50].

A. Case sampling

To answer our research questions, we conducted a comparative case study of two physical science Ph.D. programs that had redesigned their process (1) to better align their training with their perceptions of what Ph.D. graduates in their field should possess; (2) to increase transparency about their expectations; and (3) to advance diversity, equity, and/or inclusion. To select cases, we collaborated with the American Chemical Society and the American Physical Society, who invited department chairs nationally to complete a survey about the formats and structures of their Ph.D. programs' qualifying exams as well as department demographics.

We reviewed survey responses to identify Ph.D. programs for informational interviews using the following criteria: (1) programs that tracked student race and gender demographics, and (2) programs that had recently modified their transition to candidacy process. We considered programs that tracked student race and gender demographics because we believed these practices indicated that programs considered student identities when making curricular and policy decisions. The informational interviews helped us learn more about each Ph.D. program's structure and the purpose of the qualifying exam, including their motivations for changing the examination process. After completing nine informational interviews (four physics programs and five chemistry programs), we identified two chemistry Ph.D. programs and two physics Ph.D. programs to recruit as prospective case study sites. We selected these programs because of their varied transition to candidacy structures; different perceptions of the purposes of candidacy examinations; explicit focus on equity, inclusion, and diversity; and differential ranking in their respective disciplines. Of the four, one program in chemistry and one program in physics agreed to participate.[4]

Full details of the cases and their processes for advancing students to candidacy are in the findings, but we provide basic details here. The chemistry Ph.D. program was located at what we called Western University, a private, religiously affiliated, predominantly white institution that is a Carnegie-classified Doctoral University with High Research Activity. Western's chemistry Ph.D. program recently changed their transition to candidacy to create uniformity between the chemistry and biochemistry specializations. The physics Ph.D. program was located at Midwestern University, a private, predominantly white institution that is a Carnegie-classified Doctoral University with Very High Research Activity. Midwestern's physics Ph.D. program had undertaken significant structural changes to their doctoral candidacy examination and admissions processes with equity, inclusion, and diversity in mind. As is typical in the physical sciences, the composition of both departments was predominantly white and male, with racial diversity primarily coming from international doctoral students and faculty.

1. Within-case sampling

We sampled faculty, administrative staff, and students within each program to obtain multiple perspectives on learning in the transition to candidacy, both via current processes and before recent changes. We worked first with department chairs to select faculty and staff who had knowledge of the program's candidacy transition processes, and then we engaged in snowball sampling to identify additional faculty to interview. Participants identified faculty who had been involved with modifying the program's transition to candidacy process. We also collaborated with faculty, staff, and department chairs to recruit current Ph.D. students who had reached candidacy or were in the process; we similarly employed snowball sampling with them to recruit additional students. From Western, six faculty, one staff member, and two students participated; from Midwestern, four faculty, one staff member, and three students participated (Table I).

[4]Recruitment occurred as the COVID-19 quarantine was setting in; we are deeply grateful to these programs for their willingness to participate in our research during a time that was challenging for all.

TABLE II. Participant demographics by program.

		Western (chemistry)	Midwestern (physics)
Faculty	Tenured	4	2
	Men	4	1
	Women	1	3
	White	5	2
	International[a]	...	2
	Total	5	4
Staff	Women	1	1
	White	1	1
	Total	1	1
Ph.D. students	Candidates	2	2
	Men	1	2
	Women	1	1
	White	1	2
	International[a]	1	1
	Total	2	3

[a]International refers to international status for faculty or Ph.D. students.

2. Data collection

In addition to the nine informational interviews that facilitated our case selection and understanding of the focal programs' context and structures, our team conducted semistructured interviews with faculty and staff and semistructured focus groups and interviews with students in each program in Spring 2020 (Table II). Interviews "combine depth of understanding with purposeful, systematic, analytic research design to answer theoretically motivated questions" [51] (p. 159). With the onset of the COVID-19 pandemic, all data collection occurred via Zoom. Interviews were conducted by at least two team members.

Interviews with faculty ($n = 10$) and staff[5] ($n = 2$) were led by a faculty or postdoctoral research team member, with the support of a doctoral student. We inquired about participants' perceptions of the purpose of their transition to candidacy processes, the ways they helped students prepare, and the characteristics of successful qualifying exams. Interviews with staff focused on program-level messaging about the transition to candidacy and challenges that students experienced in the process. Most interviews with faculty and staff lasted 45–60 min.

Two doctoral team members led student focus groups ($n = 5$ students) and individual student interviews ($n = 2$ students), with both types of data collection procedures averaging about one hour each. Focus groups investigated students' experiences in the program and the transition to candidacy, including messaging from peers, faculty, and the program about how to navigate the process successfully. After the focus groups, we invited students to participate in individual interviews to learn more about personalized experiences in the transition to candidacy.

[5]The staff member at Western was a Graduate Program Coordinator who primarily oversaw the logistics of qualifying exam processes, including managing the paperwork. The staff member at Midwestern also managed the logistics of the qualifying exam in addition to helping faculty ensure that the qualifying exam questions were appropriate for doctoral students.

B. Analysis

All interviews were transcribed verbatim and uploaded to NVivo 12 for analysis. We engaged in multiple stages of coding before pursuing within-case and cross-case analysis. Before beginning open coding, we identified sensitizing concepts from our literature review and theoretical framework (e.g., membership, peer interactions, faculty interactions, evaluation, learning). After we completed open coding, we discussed themes within each case and codes that spanned cases. Examples of the latter included modeling, peer behaviors, (un)supportive interactions with faculty, and management of various responsibilities and tasks. Next, we engaged in axial coding, reorganizing codes and exploring relationships among them [52]. At this stage, our codebook reflected the various structural components of the transition to candidacy and the qualifying examination process.

We then developed two sets of analytic questions to interpret the data theoretically [53]. The first set of questions focused on learning as a form of movement in communities of practice at the individual and organizational levels. In analyzing the data, we asked: How are people moving toward the program's center via the transition to candidacy? Recognizing that similar dynamics of centrality and peripherality may be at play with disciplinary membership, we also designed analytic questions to explore: (1) how programs eliminated the qualifying exam and the transition to candidacy to create or protect program status in the discipline; and (2) whether and how students earned status in their programs or disciplines by earning doctoral candidacy.

Together, these analytic questions guided the construction of individual case summaries [50], which captured the exams' structure, learning objectives, and programs' motivations for transforming their exams from conventional models. Finally, we engaged in cross-case analysis by considering how patterns across the programs aligned or diverged [50] and, with themes and data from the coding process described above, addressed our research questions concerning the transition to candidacy. We offer composite narratives of each program (e.g., [54]) in the findings section and present the cross-case analysis in the discussion section.

C. Trustworthiness

We used several strategies to enhance our data and the findings' dependability, credibility, and confirmability.

Before beginning data collection, the four of us wrote separate positionality statements reflecting our experiences, beliefs, and biases about the transition to candidacy. Each of us came to this project with varied previous involvement with qualifying exams—a faculty member with ten years of experience facilitating students' transition to candidacy, a postdoc who had successfully passed their exam, and two doctoral students who had not yet taken their qualifying exams. We also reflected on how our racialized and gendered identities influenced how we conceptualized and experienced legitimation in academia.

Our research design leveraged multiple data sources to gather thick descriptions and diverse perspectives on both cases' transition to the candidacy process. We triangulated interpretations from faculty, staff, and students. We routinely wrote reflective memos throughout data collection and analysis to track how we individually and collectively made sense of the data. For instance, after each pair conducted an interview, we co-wrote a memo highlighting relevant patterns and tensions. We also routinely engaged in interrater reliability checks [49], including team discussions to talk across the cases as we constructed the narratives [55]. These activities also allowed us to be responsive to emergent understanding across the multiple rounds of analysis. Lastly, we vetted and obtained feedback on our findings with reviewers in both physical sciences and education, including one discipline-based education researcher.

D. Limitations

As noted above, recruitment and data collection during the Spring 2020 COVID-19 quarantine presented challenges. One potential limitation of our study was sampling two departments. Studying one chemistry Ph.D. program and one physics Ph.D. program allowed us to study constructions of the transition to candidacy more deeply, but including additional comparable Ph.D. programs in our sample of cases could have strengthened our cross-case comparison on the dimensions of discipline or organizational status. Another limitation may be an overrepresentation of faculty in our sample. Although we sent multiple recruitment emails to faculty, staff, and students in both departments, our final sample was unbalanced—faculty ($n = 10$), staff ($n = 2$), and student ($n = 5$) perspectives—resulting in faculty perspectives prominently informing our case narratives.

V. FINDINGS

This section discusses the process of each program in transitioning to candidacy. It also presents findings related to the stated goals and designs of exams in each program, the learning contexts and processes involved, and the cultural knowledge and behaviors that contribute to students' professional legitimacy. The programs' recognition of their position with their disciplinary status hierarchy emerged as a critical underlying factor, affecting how faculty perceived the importance of exams and candidacy, how they were willing to modify the program and candidacy transition design, and how students moved from legitimate peripheral participation toward the core of their professional communities.

A. Western university: "We're not a name-brand institution"

As a university with a midlevel ranking, Western's chemistry program viewed the qualifying exam as a means of establishing credibility within the discipline and validating its graduates as reputable scientists. Previously, students in the program's two subfields, chemistry and biochemistry, received uneven preparation for candidacy and evaluation for the qualifying exam. They had recently revamped candidacy requirements to include coursework in traditional subject matter, research design, and grant proposal writing—three areas of knowledge that professors believed would demonstrate the program's legitimacy in the disciplinary community. Each Ph.D. student also served as a lab research assistant. After students completed required coursework, including a new course designed specifically to prepare students for candidacy, students would compose and orally defend a dissertation proposal in the style of a National Science Foundation (NSF) and National Institutes of Health (NIH) proposal. The interviewed faculty indicated that the NSF and NIH writing styles were especially important.

1. Purposes of western's examinations and candidacy requirements

The purpose of assessing the proposal and how a student defended it was to link student competencies to those expected of professional chemists. A faculty member summarized this purpose as "determin[ing] if the student has what it takes to be a Ph.D. scientist." When we probed for what faculty thought this meant, they reported varied behaviors, ranging from thinking on one's feet when responding to committee questions, to clearly communicating about one's dissertation topic, to conducting independent research design, to visualizing chemistry formulas and problems. One faculty member noted paying attention to students' ability to balance lab and dissertation work. Another said that through the proposal and defense, he wanted to know if students were "technically competent and [could] think through a process." Another faculty member said that what "really distinguishes the students" came down to preparation:

> They've thought through the questions that are going to be asked, made sure that they really understand the material so that they can come in and say, "Here's what I'm presenting to you. Go

> ahead and ask me questions. I can answer pretty much anything."

Rather than being used simply to assess textbook subject matter, structural changes to the candidacy were intended to scaffold student learning about designing and communicating research.

Western's focus on scientific development was strategic. Faculty participants described the importance of their training in relation to the department's reputation and that of their graduates. "We don't admit the cream of the crop in terms of graduate students," as one stated. "So, we've gotta teach them, we've gotta bring them up to the level where they [can] get a Ph.D." The recent reforms aimed to provide students with a rigorous training that might positively influence the department's reputation:

> We're not a name-brand institution when it comes to graduate education. We have a smaller program, and so it's very important for us to have quality graduates, to be able to have a good reputation of students that we put out—ones who are prepared to be independent scientists.

Aligning the training and format of their candidacy requirements with the dissertation and NSF/NIH standards provided a mechanism for faculty and the program to convince the chemistry field that its graduates were "prepared to be independent scientists."

2. Learning the performance of scientific discourse

Faculty expected students would learn to think, write, and behave "like scientists" by engaging in an authentic task of the discipline—grant proposal development. It was in the required scientific writing course that students were introduced to the NSF/NIH proposal guidelines and how they would be expected to use these guidelines as the basis of their dissertation proposals. Field-level guidelines therefore functioned as critical tools in mediating individual student learning.

Development in culturally accepted styles of scientific writing and oral presentation was a central learning goal at Western. The scientific writing course served as a learning environment in which students could become concise, clear, and persuasive in both their written and spoken scientific communication. As a faculty member described it, the course attempted to recreate professional spaces like conferences and defenses so students could present their research to classmates and faculty and receive feedback on their performance:

> We teach in preparation for the oral presentation, and…all of our graduate students are required to take a graduate scientific writing class. Included in that writing class that I've taught before is oral presentation skills, how to give a good scientific presentation and even how to put together a good scientific poster.

Another faculty member said that the practice oral presentations taught students not to be "overly wordy, [to] have clear and concise graphs that are legible, tell the audience a story, including enough background that everybody understands, and avoid jargon." Polished presentations, precise visuals, and coherent storytelling were legitimized cultural practices in the performance of scientific discourse that was encouraged in this course.

Although a hallmark of cultural knowledge is that it is often left implicit, Western was purposeful in creating social contexts in which students would learn what faculty expected of advanced students in terms of scientific communication and self-presentation. A required course offered credit to all Ph.D. students and candidates for attending peers' dissertation and proposal defenses. First- and second-year doctoral students learned cultural nuances and expectations by observing the advanced students who were close to the program's center as a community of practice. The program deliberately designed this feature as a mechanism for transmitting disciplinary norms. One professor described advanced students as "peer mentors" who conveyed by example cultural standards about how to dress, present, and answer questions. We quote this professor at length:

> We've intentionally designed these things so that when we put these classes together, we'll put one of our more senior graduate students, somebody that has done this a couple of times, they'll be the first presenter in the semester, and they will stand up having a polished presentation, great slides, tell a great story, they'll be dressed professionally, and they'll get up and give the presentation, with the idea that the younger students that are in there are looking up at this and saying, "Oh, I thought I was just gonna be able to stand up and hand out a copy of this paper around the table and say, 'Now, let's talk about this.' " That's not gonna fly.

Presentations, they hoped, would make explicit otherwise implicit norms of performing what was construed as "polished," "great," and "professional"—in short, what would "fly" as legitimate.

During a focus group, Western students talked about peers' role in clarifying faculty expectations. One said, "You're watching that proposal presentation from other students, and you hear the questions that their committee's asking, so it's kind of helpful to know what to expect." Another student shared that hearing the doctoral candidacy committee's feedback on student oral presentations communicated an important question: "what did other people do to be successful, and how can I follow that format?" The course thus provided opportunities to observe and practice the analytical and rhetorical skills of justifying their research—critical skills for their success in the doctoral program and legitimacy as chemists.

3. Practice proposals: Learning to design legitimate projects

With guidance from faculty advisors, students first drafted their dissertation proposals in the scientific writing class. Here, faculty required students to follow NSF and NIH proposal guidelines, which channeled faculty expectations and mediated student learning. During a focus group, a student likened the dissertation proposal to a navigational tool:

> The proposal is just like a navigator, or you use Google Maps too—you have to make a plan. I use [Google Maps] to know which way to get there, how to avoid the traffic, how to have a clear way to go to your destination. So, I think the proposal gives you a clear plan to—you can get a clear mind that you know what you are doing every single day and you know your position, how far you're from your destination and how to save time, how to get help from others just to make sure you're not lost in your research.

This student believed that the dissertation proposal, like Google Maps, should provide a plan for reaching one's destination. NSF and NIH guidelines served as tools for mapping out the work that a dissertation project would require.

The dissertation proposal process, mediated by federal proposal guidelines and examples of successful proposals, involved assimilation to existing cultural standards. A student spoke of modeling their work after "accepted NIH proposal examples in my field, and kind of just read[ing] through them to get a good perspective as to what something that is successful is." A faculty member also referenced how applying for an NIH predoctoral fellowship helped one of his students comprehend and communicate research:

> One of my students did a very nice job of describing his synthesis that he was proposing of a natural product. And I think what helped him was that he had applied for an NIH predoctoral fellowship, and so he already had a lot of the material organized and had already thought about it.

In this case, the student was able to use their own past work as a standard to emulate.

Finally, the NSF and NIH proposal guidelines mediated students' learning to write persuasively, which would be applicable in any scientific career. As one faculty member noted,

> If you're an academic, you're gonna be writing a proposal. Even if you're in industry, you're probably gonna be pitching something to your boss about a project you'd like to do or an area that you'd like to move into. And it might not be as formal as an NSF 15-page limit or NIH page limit, but it's gonna have some of the same—you've gotta be persuasive in your writing. And so, this is an example, this is a way for you to practice those skills because you're gonna need them as a Ph.D. chemist.

Part of persuasiveness was feasibility. A student recalled their faculty advisor's guidance:

> Another thing [my advisor] would always say is, "Okay, does this seem feasible? Can we convince them that we're going to get this data and prove what we're saying?" So, a perspective of the proposal is like, one, the scope needs to be very tight, and then, two, how can I convince them that this is going to get done and we're halfway there?

To summarize, as a program striving to demonstrate its legitimacy in the field, Western adopted a candidacy transition process that enabled development of field-legitimated skills and styles of scientific research and communication.

B. Midwestern university: "Not everyone should be able to get a Ph.D. in physics from here"

Midwestern University was a highly regarded university, and its physics Ph.D. program was consistently ranked one of the best in the field nationally. Before its recent reform, the qualifying exam was "notoriously difficult" and a "stressful" "rite of passage" that tested knowledge in several content areas. The program replaced its high-stakes model with more accessible pathways to Ph.D. candidacy because of gender disparities in pass-fail rates. Under the new model, students were still required to demonstrate knowledge in classical mechanics, electrodynamics, quantum mechanics, and statistical mechanics; however, they could do so by passing low-stakes graduate diagnostic exams (GDEs) before starting coursework, by receiving at least a D grade in core courses, or by completing a combination of the two. These changes have mitigated concerns about gender equity and student well-being. Although the previous exam structure did not scaffold student learning about how to become scientists, some participants still worried that the new system may be leaving too much essential cultural learning about becoming a scientist to the informal—and therefore potentially inequitable—processes that take place in labs after students have achieved candidacy.

Prestige was a key quality of the Ph.D. program's organizational culture. It shaped members' beliefs that a Ph.D. from their program should not be easy to obtain and that their extremely selective admissions afforded possibilities for relaxing expectations around candidacy. One faculty member shared, "I think it's fair to say that not everyone should be able to get a Ph.D. in physics from Midwestern…it's not the case that anyone who wants to can do it." This mindset was tempered by emergent understandings of how the Ph.D. admissions process was a gatekeeping tool that had preselected students for desired skills and qualities that would serve them as physicists. Participants assumed that if admitted students were deemed "good enough" for their elite program, they could reach candidacy. Under this holistic admissions process, "good enough" meant it was not purely a matter of textbook knowledge. An administrative staff member said they also considered "how well did you do in the challenges that you did face" and "an ability to technically evaluate literature and be able to do technical troubleshooting."

In short, the department completely eliminated the qualifying exam because it believed with confidence that its students were already "Midwestern material"—that is, they possessed some qualities of legitimated physicists. One student shared,

> This was already a hard program to get into, and I think the reason they cut the super hard candidacy exam was that they were letting most kids through, but it was just an enormous psychological cost. Why have two bottlenecks where the second one [i.e., the former qualifying exam] isn't a good bottleneck but it's extremely stressful?

This student also went on to characterize the qualifying exam as "totally unnecessary" because of the competitiveness of admissions and its selection for qualities like initiative that, in some universities, was developed by doctoral education or the candidacy process. In the next section, we describe how learning the practice of self-sufficiency may have come at the cost of opportunities for learning other important cultural knowledge through coursework and research lab advising.

1. Learning an unspoken cultural practice of self-sufficiency

As we interviewed faculty and students, it became evident that an implicit learning outcome in Midwestern's program was developing self-sufficiency. Indeed, the expectation that students would develop in this way may help explain why the program selected for initiative during the admissions process. One faculty member spoke at length about doctoral education as a time for students to take initiative and learn independence:

> You don't go to graduate school to do coursework, right? So, I mean it's something you do. You learn some topics, but basically graduate school is a time when you learn how you can learn things *by yourself*…. The utility of coursework is that it opens you up to other concepts that you didn't learn as an undergraduate. But I mean coursework is appropriate for a master's-level degree, right? But for the doctorate level, I don't think the coursework has anything to do to prepare you to be a professional scientist.

The faculty member continued to emphasize that students learned to be "resourceful," take "ownership of a project," and "control what they work on"—dimensions of self-sufficiency and necessary qualities for physicists.

Although removing the written and oral qualifying exams eliminated a hurdle, it also left a void where once they had assessed students' readiness for lab and dissertation research. "There's no assessment that's occurring at that stage," one professor put it. "Right now, the assessment is whether a student can find a research group successfully…. There's no exam. It's just if the student finds a thesis advisor." Having designed a program where students did not enter a lab until after candidacy and were not assigned a thesis advisor, Midwestern faculty came to rely on whether students could successfully navigate these new hurdles as a proxy for the self-sufficiency they would need for independent research.

2. Missed opportunities for professional learning

We saw in the case of Western a clear example of scientific socialization taking place within required, credit-bearing coursework. Coursework did not operate this way in Midwestern's Ph.D. program. Students experienced "candidacy courses" as "an extension of undergrad" rather than a meaningful learning context. With courses focusing more on breadth than depth of content, students felt as if they were reviewing concepts they had already learned—just at a faster pace and more superficially. Student focus group participants expressed frustration that precandidacy coursework often felt like "just another year of undergrad." Instead, faculty and students highlighted research labs as consequential learning spaces that helped students to move closer to the center of the community of practice in both the department and discipline. As one student said, "I came to graduate school to do research, and that's something that I haven't had much time for as of late just because of the candidacy courses."

One course included an opportunity to engage in research, but students portrayed it as a missed opportunity for learning that reinforced Midwestern's unspoken norm of self-sufficiency. The course required each student to design and carry out an experimental project and provided research lab rotations to experimentalists and those interested in applied physics. One student in a focus group described her course project in depth:

> We start with an advisor. Generally speaking, there should be some hardware involved, so you have to, as part of your project, learn some technical skills.… There's lots of different types of, you know, technical knowledge that you could acquire, and also you have to compile your data and your results into a presentable format and present that in both a presentation, a paper, and a poster session. Just being able to analyze your data in a way that's useful and present it in a way that you communicate your science to people.

The student described how the project served as a site where expectations for managing a lab were clarified and where she could practice various research skills, but without feedback, her learning was hindered. When we asked the focus group about the evaluation of these projects, another student laughed and bluntly responded, "Nobody knows…there has been no feedback." A third noted, "All of the pressure I received was from the graduate student I was working with on that project and the professor who really cared what sort of stuff our lab puts out."

Both faculty and students explained that an essential part of becoming a physicist was learning to solve problems, while opportunities for problem solving in lab settings are radically different from those available in classroom contexts. One faculty shared that "many students feel competent in physics because they are good at problem solving…and physics problems always have clear-cut answers, [but] research doesn't." She further described how even her "best physics students" had difficulty translating their "technical capabilities" and problem-solving skills from coursework to applied research. Although the experimental course project presented potential for cultural learning, and although it involved the assignment of meaningful tasks, it could have been more beneficial with enhanced instructor engagement. It might not be that students at Midwestern need the same scaffolding in project development as those at Western, but instructors could purposefully provide support for translating their technical capabilities to the research context. This missed opportunity is especially noteworthy, given recent findings that instructor-student interactions in lab-based physics courses strongly correlate to the formation of self-efficacy [44].

In summary, Midwestern's decision to replace the qualifying exam with GDEs and core coursework reduced anxiety around achieving candidacy. However, the new program structure's dual pathways did not address the issues of limited advising and unequal research opportunities early in the program. Although this is unfortunate for students' access to opportunities for learning scientific legitimacy early in their training, it is fitting that in such a prestigious program, students were selected for and experienced anticipatory socialization to the broader disciplinary norm of self-sufficiency. This too, one could argue, is part of learning legitimacy in physics.

VI. DISCUSSION

We applied concepts from sociocultural theory to investigate (1) how faculty and students constructed the purposes of the qualifying exam and transition to candidacy, and (2) how Ph.D. programs facilitated doctoral student learning in association with that transition. Our findings highlighted how qualifying exams and the transition to candidacy can move students toward the center of scholarly communities of practice through formal and informal learning that socializes them to disciplinary norms (e.g., self-sufficiency) and legitimizes them through the acquisition of expected knowledge and skills (e.g., giving polished presentations, writing well-constructed proposals). Such movement happens at two levels: organizational and individual. This section begins with a cross-case comparison before turning to research and practice implications.

A. Cross-case comparison

Both cases represent redesign efforts that attend to student well-being but vary in developing students as emerging knowledge creators and stewards of the discipline [37]. Our findings align with those of Anderson *et al.* [13] and Estrem and Lucas [12], who found that Ph.D. programs demonstrate their legitimacy through rigor in training and requirements for candidacy.

1. Status considerations

Faculty perceptions of their program's disciplinary status (i.e., "not a name brand" vs "not everyone deserves a degree from here") shaped how they understood and justified changes to candidacy requirements and program structures. Faculty in both programs referred to their program's admissions selectivity as a factor that guided redesign options. This is a small but important point about how systemic change in doctoral education works in practice: Western felt a sense of urgency around professional development because they knew that their admissions did not yield the "best" recruits. Faculty at Western therefore made efforts to develop students' skills and confidence in activities like communicating research, which would facilitate their success in future federal grant competitions. Elite status and a highly selective admissions process gave Midwestern the confidence to eliminate the qualifying exam entirely. Professors there did not scaffold students in how to access research labs, match with thesis advisors, or compete for research opportunities. Professional development was not part of the formal curriculum.

2. Learning legitimacy in communities of practice

The two cases also revealed differences in the potential of required doctoral coursework to support the process of "becoming a scientist." Both Ph.D. programs purposefully used multiple contexts to support and mediate learning. Consistent with sociocultural theories [33,34], we found that learning for candidacy is "situated learning" because it occurs in specific courses and research environments; in programmatic structures that define course curriculum and Ph.D. requirements; and in wider discipline-level conversations driving recent changes to exam structures. The difference was this: Western's beliefs about its students' preparedness, or lack thereof, led its faculty to be much more explicit about using coursework to develop students' professional behaviors and norms. Midwestern assumed that students would learn professional development skills in the lab.

At Western, the program used federal agencies' proposal guidelines as a framework for the dissertation proposal and the capstone of a required scientific writing course, in which it was hoped students would learn and practice the performance of scientific discourse. Learning to be concise, clear, and persuasive in written and spoken scientific language was believed to be the key competency worthy of formal development. Proposal guidelines thus served as cultural tools that clarified field norms. Western also awarded course credit for attending advanced students' proposal and dissertation defenses, which participants cited as settings that transmitted local expectations for effective scientific communication. Here, becoming a scientist was an explicit process: faculty prepared students for known competition for NSF/NIH resources, whose winners would be defined as legitimate disciplinary actors. In training first- and second-year students to behave and think like scientists generally, and write proposals like funded chemists specifically, faculty hoped the program would move closer to the center of the discipline.

The transition to candidacy at Midwestern had less to do with scaffolding student movement into legitimate disciplinary membership, although it did socialize them to the absence of support—and necessity of self-sufficiency—in the physics community of practice. Students were understood as moving closer to the center of the discipline when they learned how to translate technical and problem-solving skills developed in the classroom to specific research problems. Other than a cursory introduction to these matters in the lab rotation course (for which students reported receiving insufficient feedback), formal opportunities for hands-on scientific socialization and legitimation were few before candidacy. Instead, students passed required coursework or graduate diagnostic exams before being allowed to do what was experienced as the "real work" of research. Unlike Western, where coursework mediated scientific development, students and faculty alike at Midwestern perceived coursework as a necessary hurdle to clear before entering the lab, where socialization would really begin.

B. Implications for research

In what follows, we offer implications for research on informal learning contexts and racialization. We also encourage scholars to examine the same research questions in departments that have not recently reformed their qualifying exams and/or those that do not presently prioritize equity and inclusion in their curriculum and policy decisions. It is possible that the significance and purpose of qualifying exams would vary because of differing values and goals. Given our findings, these examinations are important particularly for considering how to foster change in support of student learning and equity.

1. Informal learning contexts and opportunities

Participants in both Ph.D. programs highlighted the power of informal learning contexts and mechanisms (e.g., research labs, peer relationships, practice presentations, *ad hoc* mentoring). In both cases, there was a consensus that "real" learning could not be limited to the classroom. Investigating both formal curriculum and informal learning contexts in future studies of graduate education would offer powerful insights into what and how students learn; what resources contribute to learning; and which structures and practices inhibit success and equity as students transition from student to candidate and beyond in the postdoctoral years.

Learning in informal and extracurricular contexts has significant implications for equity that are worthy of exploration. Faculty may disproportionately afford access to spaces that provide informal but crucial support to students who have dominant identities or align with dominant images of scientists, furthering racial and gender exclusion. When access to mentoring relationships, research opportunities, and study groups are left to informal social dynamics, it often reflects and can reinforce segregated social networks [56]. Researchers should examine how faculty and advanced students welcome early career graduate students into their networks and labs (or choose not to), specifically focusing on which qualities and individuals they deem valuable.

2. Examining embedded racialization

In this paper, we began to examine how organizational cultural norms like status and legitimacy inform the way faculty, staff, and students construct the purposes of qualifying exams and which options they have for changing the requirements for transitions to candidacy. These and other sites for the performance of scientific discourse are opportunities for researchers to address unexamined racialization, given the biases known to plague (un)structured evaluation contexts (see, for example, [57]). To illustrate, performances of scientific communication that privilege unwavering conformity to norms of standard English may disproportionately hurt the advancement of rising ethnically and racially minoritized scholars. Our primarily white participants were ill-equipped to reflect on racialization in the embodiment and performance of “good science.” Research on doctoral education should not take for granted how disciplines—and the bodies of knowledge and the origins of intellectual work produced within disciplines—often reflect racialized and gendered norms that privilege whiteness and masculinity [36]. Researchers should explore the experiences of Black, Indigenous, and other racially minoritized early-career scholars to illuminate and interrogate racialized and gendered practices and how they inhibit learning and validation as scholars.

C. Implications for practice

Reimagining doctoral education as a process of cultivating disciplinary membership, versus solely demonstrating readiness for dissertation research, comes with distinct design opportunities. When the core purpose is for faculty to evaluate doctoral students’ progression toward mastery of accepted socioculturally constructed knowledge and methods, then a standardized format and content are logical [30]. However, this posture is predicated on students needing to prove themselves, a potentially unnecessary assumption. Further, standardization leaves doctoral students with minimal agency in selecting the areas in which they will focus their learning or be assessed, including research specialty [2]. Candidacy transition processes that include the development and review of a dissertation proposal, and perhaps a dissertation committee, send a message of belonging and develop student skills in content areas with relevance to their developing expertise. We offer recommendations to help Ph.D. faculty and staff make structural changes with attention to underlying cultural values.

1. Design for intended learning outcomes

Changing the tasks that students are required to perform without asking how and why students are expected to learn them may fail to reduce stress, increase student success, or improve alignment of activities with Ph.D. outcomes. Rather than defining only what they want to avoid (e.g., high-stakes exams), doctoral programs should design learning around specific outcomes, acknowledging their motives for those aims. Faculty and staff should first explore their assumptions about what students should know and do. After identifying desired outcomes, we encourage intentional planning toward these goals. Scholarship on intentional design for learning outcomes highlights the interconnected nature of learning objectives, assessments, and strategies [58]. A well-designed, transparent system of training should include objectives that state what students should be able to demonstrate (e.g., students will construct an independent research proposal according to NSF guidelines) as well as *how* and *where* these skills are fostered and assessed.

Transparency means being clear about what is expected to learn and the rationales for where, how, and why students will do so [59]. This focus on transparency should be used to review and (re)design transition to candidacy processes (e.g., preparatory courses, exams, defenses). For example, self-sufficiency is an unspoken norm in Midwestern’s program and was constructed as an important skill for students to succeed as physicists. If the university wanted to institutionalize this as a learning goal, it could identify *where* and *how* it assesses students’ development of self-sufficiency, *where* and *how* students learn that this is a necessary skill, and *where* and *how* students develop the skill itself.

VII. CONCLUSION

Through this comparative case study, we hope to open new avenues for studying doctoral education and motivate Ph.D. program leaders to create more inclusive and developmentally oriented candidacy transitions. The Western case study illustrates how programs can reconceptualize candidacy requirements in terms of learning scientific community norms, especially regarding communication via presentations and proposals. There, training and candidacy requirements were intertwined to legitimate students as scientists. This interwoven training also illustrates how to make explicit disciplinary cultural practices that are usually left implicit. Although Midwestern transformed candidacy and qualifying exam processes, its story is less about students learning the field’s legitimated cultural practices. Midwestern’s new structure, if anything, recentered conventional subject matter and has yielded cautions about waiting until candidacy to expose students to scholarly research. A full story of learning legitimacy in that program would require a study of lab-based learning. Midwestern nonetheless exemplifies how even “elite” programs can assess content knowledge while eliminating harmful, high-stakes exams. We hope that researchers and practitioners will continue to center student learning and success in examining and refining doctoral education across the disciplines.

ACKNOWLEDGMENTS

This work was funded by National Science Foundation Grant No. 1834528; the opinions stated in this paper are our own.

[1] V. I. Baker and M. J. Pifer, The role of relationships in the transition from doctoral student to independent scholar, Studies Cont. Educ. **33**, 5 (2011).
[2] A. L. Furstenberg and A. Nichols-Casebolt, Hurdle or building block, J. Teach. Social Work **21**, 19 (2001).
[3] J. Harshman, Review of the challenges that face doctoral education in chemistry, J. Chem. Educ. **98**, 259 (2021).
[4] American Center for Physics, Graduate education in physics: Which way forward?, *Proceedings of a Conference to Discuss the Status and Future of Graduate Education in Physics, College Park, MD* (American Physical Society and American Association of Physics Teachers One Physics Life, College Park, 2008).
[5] J. C. DiPietro, W. Drexler, K. Kennedy, V. Buraphadeja, F. Liu, and K. Dawson, Using Wikis to collaboratively prepare for qualifying examinations, TechTrends **54**, 25 (2010).
[6] L. Owens, B. M. Zwickl, S. V. Franklin, and C. W. Miller, Misaligned visions for improving graduate diversity: Student characteristics vs. systemic/cultural factors, presented at PER Conf. 2018, Washington, DC, 10.1119/perc.2018.pr.Owens.
[7] J. E. McLaughlin, K. Morbitzer, M. Meilhac, N. Poupart, R. L. Layton, and M. B. Jarstfer, Standards needed? An exploration of qualifying exams from a literature review and website analysis of university-wide policies, Stud. Grad. Postdr. Educ. (to be published), 10.1108/SGPE-11-2022-0073.
[8] National Research Council, *Challenges in Chemistry Graduate Education: A Workshop Summary* (The National Academies Press, Washington, DC, 2012).
[9] R. S. Barthelemy, M. Lenz, A. Knaub, J. Gerton, and P. Sandick, Graduate program reform in one department of physics and astronomy: From tragedy to more progressive policies and an evolving culture, Phys. Rev. Phys. Educ. Res. **19**, 010102 (2023).
[10] J. E. Nardo, Case for narrative inquiry: Understanding the experiences of Elena, a historically minoritized graduate student in chemistry, J. Chem. Educ. **99**, 113 (2022).
[11] I. Esmonde, Power, and sociocultural theories of learning, *Power, and Privilege in the Learning Sciences: Critical, and Sociocultural Theories of Learning*, edited by I. Esmonde and A. N. Booker (Routledge, New York, 2017), pp. 6–27.
[12] H. Estrem and B. E. Lucas, Embedded traditions, uneven reform: The place of the comprehensive exam in composition and rhetoric Ph.D. programs, Rhetoric Rev. **22**, 396 (2003).
[13] W. P. Anderson, C. J. Krauskopf, M. E. Rogers, and G. W. Neal, Reasons for comprehensive examinations: A reevaluation, Teaching Psychol. **11**, 78 (1984).
[14] K. Kostohryz, The doctoral comprehensive examination in counselor education: Faculty members' perception of its purposes, J. Counselor Preparation Supervision **8**, 1 (2016).
[15] D. Campbell, T. Appelquist, R. Diehl, J. Fajans, J. D. Garcia, J. Gates, A. Goldman, P. Jung, and M. Paesler, *Graduate Education in Physics*, Report of the Task Force on Graduate Education in Physics (American Association of Physics Teachers and American Physical Society, College Park, MD, 2006).
[16] J. Posselt and R. Liera, *Doctoral Candidacy and Qualifying Exams: A Guide to Contexts, Costs for Equity, and Possibilities of Transformation* (Equity in Graduate Education Resource Center, 2023), https://equitygraded.org/rsrc/qualifyingexams/.
[17] E. A. Schmidt, L. E. Homeyer, and J. L. Walker, Predictors of success on the counselor preparation comprehensive examination, Couns. Educ. Superv. **48**, 226 (2009).
[18] C. R. McAdams, D. L. Robertson, and V. A. Foster, Doctoral oral examinations and contemporary counselor education: Are they compatible?, Couns. Educ. Superv. **52**, 270 (2012).
[19] N. Ponder, S. E. Beatty, and W. Foxx, Doctoral comprehensive exams in marketing: Current practices and emerging perspectives, J. Market. Educ. **26**, 226 (2004).
[20] S. M. North, *Refiguring the Ph.D. in English Studies: Writing, Doctoral Education, and the Fusion-based Curriculum* (National Council of Teachers of English, 2000).
[21] J. Posselt, R. Liera, A. J. Rodgers, and L. N. Irwin, Rethinking the race, *Inside Higher Education* (2021), https://www.insidehighered.com/advice/2021/05/04/need-reconsider-qualifying-exams-phd-candidacy-opinion.
[22] S. Gelmez Burakgai and A. Yildirim, Journey into doctoral candidacy: A grounded theory of doctoral qualification exam process, Croatian J. Educ. **19**, 63 (2017).
[23] J. R. Posselt, *Inside Graduate Admissions: Merit, Diversity, and Faculty Gatekeeping* (Harvard University Press, Cambridge, Massachusetts, 2016).
[24] K. Rosinger, K. Ford, J. Posselt, and J. Choi, Exploring the impact of GRE-accepting admissions on law school diversity and selectivity, Rev. High. Educ. **46**, 109 (2022).
[25] R. Liera and T. E. Hernandez, Color-evasive racism in the final stage of faculty searches: Examining search committee hiring practices that jeopardize racial equity policy, Rev. High. Educ. **45**, 181 (2021).
[26] D. K. White-Lewis, The facade of fit in faculty search processes, J. Higher Educ. **91**, 833 (2020).
[27] M. Lamont, *How Professors Think: Inside the Curious World of Academic Judgment* (Harvard University Press, Cambridge, Massachusetts, 2012).
[28] J. L. Riviere, Qualified…for what? The potential for qualifying exams to become meaningful professional milestones, *Emerging Directions in Doctoral Education* (Emerald Group Publishing Limited, Bingley, 2016), Vol. 6, pp. 243–258.
[29] E. Marshman, R. Sayer, C. Henderson, and C. Singh, Contrasting grading approaches in introductory physics and quantum mechanics: The case of graduate teaching assistants, Phys. Rev. Phys. Educ. Res. **13**, 010120 (2017).
[30] A. H. Herzig, Where have all the students gone? Participation of doctoral students in authentic mathematical activity as a necessary condition for persistence toward the Ph.D., Educ. Studies Math. **50**, 177 (2002).
[31] A. J. MacLachlan, Preservation of educational inequality in doctoral education: Tacit knowledge, implicit bias and university faculty, UC Berkeley Center for Studies in Higher Education, 2017.

[32] S. L. Pressey, L. C. Pressey, and E. J. Barnes. The final ordeal, J. Higher Educ. **3**, 261 (1932).

[33] J. Lave and E. Wenger, *Situated Learning: Legitimate Peripheral Participation* (Cambridge University Press, Cambridge, England, 1991).

[34] B. Rogoff, *The Cultural Nature of Human Development* (Oxford University Press, New York, 2003).

[35] R. S. Barthelemy, M. McCormick, C. R. Henderson, and A. Knaub, Educational supports and career goals of five women in a graduate astronomy program, Phys. Rev. Phys. Educ. Res. **16**, 010119 (2020).

[36] A. D. Robertson, V. Vélez, W. T. Hairston, and E. Bonilla-Silva, Race-evasive frames in physics and physics education: Results from an interview study, Phys. Rev. Phys. Educ. Res. **19**, 010115 (2023).

[37] G. E. Walker, C. M. Golde, L. Jones, A. C. Bueschel, and P. Hutchings, *The Formation of Scholars: Rethinking Doctoral Education for the Twenty-First Century* (Jossey-Bass, San Francisco, 2008).

[38] L. D. Gonzales and A. L. Terosky, From the faculty perspective: Defining, earning, and maintaining legitimacy across academia, Teach. Coll. Rec. **118**, 1 (2016).

[39] M. C. Suchman, Managing legitimacy: Strategic and institutional approaches, Acad. Manag. Rev. **20**, 571 (1995).

[40] E. W. Schoon, Operationalizing legitimacy, Am. Sociological Rev. **87**, 478 (2022).

[41] J. C. Weidman, D. J. Twale, and E. L. Stein, *Socialization of graduate and professional students in higher education: A perilous passage?* (ASHE-ERIC Higher Education Report, San Francisco, CA, 2001).

[42] T. Becher and P. Trowler, *Academic Tribes and Territories*, 2nd ed. (The Society for Research into Higher Education & Open University Press, New York, 2001).

[43] E. Wenger, Communities of practice and social learning systems: The career of a concept, in *Social Learning Systems and Communities of Practice*, edited by C. Blackmore (Springer, Cham, 2010), pp. 179–198.

[44] D. J. Rosen and A. M. Kelly, Working together or alone, near, or far: Social connections and communities of practice in in-person and remote physics laboratories, Phys. Rev. Phys. Educ. Res. **18** (2022).

[45] K. Holley, Animal research practices and doctoral student identity development in a scientific community, Studies Higher Educ. **34**, 577 (2009).

[46] K. A. Holley, Doctoral education and the development of an interdisciplinary identity, Innov. Educ. Teach. Intl. **52**, 642 (2015).

[47] D. J. Phillips and E. W. Zuckerman, Middle-status conformity: Theoretical restatement and empirical demonstration in two markets, Am. J. Sociol. **107**, 379 (2002).

[48] S. B. Merriam and E. J. Tisdell, *Qualitative Research: A Guide to Design and Implementation*, 4th ed. (Jossey-Bass, San Francisco, 2016).

[49] R. K. Yin, *Case Study Research: Design and Methods*, 5th ed. (Sage, Newbury Park, California, 2014).

[50] R. E. Stake, Qualitative case studies, in *Handbook of Qualitative Research*, edited by N. K. Denzin and Y. S. Lincoln (Sage, Newbury Park, California, 2005), pp. 443–462.

[51] M. Lamont and A. Swidler, Methodological pluralism and the possibilities and limits of interviewing, Qual. Sociol. **37**, 153 (2014).

[52] J. Saldaña, *The Coding Manual for Qualitative Researchers* (Sage, Newbury Park, California, 2016).

[53] A. Neumann and A. M. Pallas, Critical policy analysis, the craft of qualitative research, and analysis of data on the Texas top 10% law, in *Critical Approaches to the Study of Higher Education: A Practical Introduction*, edited by A. M. Martínez-Alemán, B. Pusser, and E. M. Bensimon (Johns Hopkins University, Baltimore, Maryland, 2015), pp. 153–173.

[54] J. Posselt, K. B. Porter, and A. Kamimura, Organizational pathways toward gender equity in doctoral education: Chemistry and civil engineering compared, Am. J. Educ. **124**, 383 (2018).

[55] S. R. Jones, V. Torres, and J. Arminio, *Negotiating the Complexities of Qualitative Research in Higher Education: Fundamental Elements and Issues*, 2nd ed. (Routledge, New York, 2014).

[56] M. Ong, C. Wright, L. Espinosa, and G. Orfield, Inside the double bind: A synthesis of empirical research on undergraduate and graduate women of color in science, technology, engineering, and mathematics, Harv. Educ. Rev. **81**, 172 (2011).

[57] D. Kahneman, *Thinking, Fast and Slow* (Macmillan, New York, 2011).

[58] Eberly Center, *Design and Teach a Course: Articulate Your Learning Objectives* (Eberly Center for Teaching Excellence & Educational Innovation, 2022).

[59] M. A. Winkelmes, M. Bernacki, J. Butler, M. Zochowski, J. Golanics, and K. H. Weevil, A teaching intervention that increases underserved college students' success, Peer Rev. **18**, 31 (2016).