Investigating faculty perspectives on written qualifying exams in physics

Shiva Basir [1] and Eric Burkholder [1,2,*]

[1] *Department of Physics, Auburn University, Auburn, Alabama 36849, USA*
[2] *Dpeartment of Chemical Engineering, Auburn University, Auburn, Alabama 36849, USA*

(Received 28 November 2023; accepted 2 April 2024; published 10 May 2024)

Doctoral qualifying exams are considered essential in assessing a student's readiness for research and advanced studies. Despite their significant role in many physics programs, questions have been raised about their format, execution, and relevance. Our research investigates perceptions held by faculty members regarding the graduate doctoral examination (GDE), a written qualifying exam in Auburn University's physics department doctoral program. We used a combination of semistructured interviews and a survey to probe their viewpoints about the purpose and necessity of written qualifying exams, their role in student preparation for these exams, and the efficacy of these exams in measuring students' comprehensive knowledge and potential for success in physics. Despite the general consensus on the necessity of the GDE, faculty members expressed doubts about its ability to accurately predict students' future research success and its alignment with other graduate program elements such as coursework. Proposed modifications ranged from an emphasis on oral assessments and research presentations to a complete overhaul of the examination structure. Despite these suggestions for change, the lack of agreement on a specific alternative underscores the complexity of executing substantial modifications to the GDE. Our study contributes to the ongoing dialogue on optimizing doctoral qualifying exams to better serve students and academic institutions.

DOI: 10.1103/PhysRevPhysEducRes.20.010139

I. INTRODUCTION

Doctoral qualifying exams, often serving as the gateway to candidacy status and independent research, have been a mainstay in higher education. Doctoral qualifying exams are designed to assess the depth and breadth of a student's knowledge and verify their capabilities for independent research [1,2]. Depending on the discipline, the exams can be subject-specific tests probing content knowledge or can be designed to ensure a general understanding of the research process. A study by Gardner investigated the variations in qualifying exams across disciplines and countries [3]. For science, technology, engineering, and mathematics (STEM) disciplines, the exams often include written and oral components focusing on the proposed research. In contrast, humanities and social sciences typically involve comprehensive essays on various topics, followed by an oral defense [4]. This variety reflects the unique demands and practices of each discipline. However, there are controversies surrounding these exams' effectiveness and validity, for example, because of their considerable contributions to departures from graduate programs [5,6]. In this study, we investigated how physics faculty think about written qualifying exams, as there is a consensus in physics that such exams are a necessary part of a doctoral education [2].

We sought to explore general themes of why we use qualifying exams, how the process is carried out, and then what we achieve through this process. For this study, we limit ourselves to the perspectives of faculty members, as they are the ones who maintain and administer these exams. Our research questions for this study were as follows:

1. What are faculty members' perceptions regarding the purpose and necessity of written qualifying exams?
2. How do faculty members perceive their role in preparing students for qualifying exams and what are their views on the relationship between such qualifying exams and the coursework in a graduate program?
3. How do faculty members perceive the effectiveness of qualifying exams in measuring students' knowledge, skills, and potential success in physics?
4. What do faculty members suggest as modifications to or alternatives for qualifying exams?

A. Background

Qualifying exams emerged in the 20th century as an academic mechanism designed to ensure Ph.D. candidates'

*Corresponding author: ewb0026@auburn.edu

competency and preparedness before embarking on research [7]. The exams were designed to act as a quality control measure, targeting students' understanding of key theories, topics, and methodologies in the discipline [8,9]. For many academic departments, these evaluations represent a blend of tradition and standard setting, simultaneously acting as a badge of "rigorous" academic training and a filtering mechanism for doctoral candidates [7,10]. However, some have started to question the idea that these exams reflect a certain amount of academic "rigor." These exams can become sources of stress and anxiety for students, especially when perceived as rites of passage rather than genuine measures of capability [2,11]. Moreover, the assumption that such exams, in their traditional format, accurately evaluate a student's potential has also come under scrutiny [12,13]. The design of qualifying exams often faces criticisms due to arbitrary decision-making processes regarding essential topics and format [14,15]. Evaluations can be perceived as nontransparent, often leading to confusion among candidates about expectations and grading metrics [14,16,17], and concerns about implicit bias [18].

Pressey and colleagues identified criticism of qualifying exams' inconsistent evaluation methods and unclear standards as early as the 1930s [19]. Current literature still echoes these concerns, suggesting that despite the evolution in academic landscapes, the inherent issues surrounding qualifying exams persist [20]. Recent calls in the literature emphasize the importance of infusing sociocultural perspectives and understanding cultural norms that shape perceptions of academic "readiness" [20]. A reimagined vision for doctoral education that is holistic and student centric has also been advocated.

Central to the discourse on qualifying exams is the idea of "student deficit." Student deficit models of learning emphasize student shortcomings or failures to meet set standards, often overlooking systemic or structural barriers that might play a role [21]. This deficit-based perspective could inadvertently shape the design and execution of qualifying exams, potentially perpetuating inequalities. The concept of student deficit primarily centers on the notion that students, especially from marginalized or underprivileged backgrounds, are perceived as inherently lacking—be it skills, knowledge, or cultural capital. The notion that qualifying exams measure a student's competency and preparedness for research reflects a deficit mindset, implying that students are not "good enough" to embark on research until they have proven otherwise.

It is worth noting that, except for Ref. [2], most of the research on the history and impact of qualifying exams comes from other disciplines. While the general themes of "readiness for research" and maintaining "academic rigor" are likely common across all disciplines, including physics, there may be certain aspects of these exams that are more unique to physics or even individual departments. For example, the core content is likely much more standardized across physics programs than it may be in other disciplines, as much of what is considered fundamental physics knowledge has changed relatively little in the past 70 years. Similarly, the format of qualifying exams is likely more standard across physics departments, as physics tends to place emphasis on mathematical calculations and view calculations as objective truths [22]. Also, physics has a unique disciplinary culture that elevates models of the "lone genius," and concepts of individuality that may make discourse around the change of qualifying exams more difficult [23].

B. Institutional context

Auburn University's Physics Department doctoral program offers a standard course structure in the first year, with a written qualifying exam called the graduate doctoral examination (GDE) at the end of the first summer term. The program structure is summarized in Table I. During the first two semesters, students enroll in four core courses: classical mechanics, statistical mechanics, electromagnetism, and quantum mechanics. Each of these courses carries three credits. Along with these core courses, students also select two elective courses and take a one-credit pedagogy class and a one-credit research seminar. Students must maintain a minimum 3.0 GPA to remain in good academic standing. First-year students are also required to serve as teaching assistants (TAs) for two sections of introductory lab courses in the fall and three sections in the spring

TABLE I. Structure of first-year coursework Auburn University's physics doctoral program.

Semester	Courses	Credits
Fall and Spring	Core courses (classical mechanics, statistical mechanics, electromagnetism, quantum mechanics)	12 (3 each)
Fall and Spring	Elective courses	Variable
Fall and Spring	Pedagogy class	1
Fall and Spring	Research seminar	1
Summer	Graduate doctoral examination (GDE)	⋯
	Total credits	18+

semester. This TA responsibility includes holding office hours for three hours a week, grading assignments, and proctoring exams. Serving as a TA covers students' tuition and provides a stipend to cover living expenses. It is also worth noting the existence of other "hidden" responsibilities of physics graduate students such as attending seminars, trying to find research advisors, etc.

The GDE is administered at the end of the summer term following students' first year in the program. The GDE is a set of four comprehensive exams covering the four core content areas and typically spans 2 weeks, with each exam lasting 4 h. Students must score at least 50% on each exam to pass the GDE, and students who do not meet this passing threshold are offered a second opportunity to take the exam the following spring. Students who do not pass all four exams by the second attempt are dismissed from the program and are typically offered the option to obtain a nonthesis Master's degree. Students do not start their research until they have completed at least their first semester, though this is often delayed until they pass the GDE. This is not an official policy, but it is strongly recommended to students that they focus on their coursework and GDE preparation before getting involved in research.

II. METHODOLOGY

We conducted a research study on faculty perspectives of the graduate doctoral examination (GDE) in the Physics Department at Auburn University in 2022–2023. We interviewed eight faculty members from various subfields and ranks and then conducted a survey of the entire faculty (24 members). Due to the exploratory nature of our research questions, we employed a grounded theory approach to analyzing the data. Grounded theory is a research methodology that aids the discovery of emergent patterns and structures within data [24]. In contrast to other methodologies that test a preexisting theory, grounded theory allows for the generation of new theories and concepts grounded in the data itself [25].

Our interpretive lens for this study was interpretive phenomenological analysis (IPA). IPA is an approach that enables researchers to deeply explore and understand how individuals perceive, interpret, and make sense of their personal and social world. In this context, it was used to delve into how individuals within the Physics Department perceive and interpret the GDE and its influences on their work. The IPA encouraged empathetic engagement with participants' lived experiences, granting rich and nuanced insights into their personal perceptions of the GDE [26–28]. IPA played an important role in shaping the interview protocol, and the survey to some extent, by explicitly talking about the role of the GDE in these faculty members' lives, either in their experiences as a graduate student or their experiences facilitating an exam that they may not agree with. We also allowed the interviewees to articulate their vision for the graduate program and how the GDE would or would not fit into that ideal. IPA also played a role in our interpretations of the data, particularly by acknowledging how some faculty members' experiences may have shaped their current perspectives.

Our interpretive standpoint is based on the premise that a detailed exploration of the subjective experiences and perceptions of individuals within the department can offer profound insights into departmental practices [29,30]. We have employed a combination of one-on-one interviews and comprehensive surveys as our data gathering tools. Initial in-depth interviews were conducted with eight faculty members, providing us with rich insights into their personal experiences, perceptions, and attitudes related to the GDE. Informed by the insights garnered from these interviews, we crafted and conducted a survey, extending the invitation to all faculty members within the department. The response rate was 96%. This broad participation enabled us to identify trends and highlight commonalities and differences.

A. Confidentiality

Maintaining participant anonymity was a critical aspect of this research, although it presented challenges due to the small size of the Physics department at Auburn University. E. B. took precautions to anonymize the interview data before providing access to S. B.; the survey data were completely anonymous and the only identifying information collected was whether a faculty member was tenure-track or non-tenure track, and whether they were a theoretician or experimentalist. Participants' responses were coded and securely stored in a password-protected Box folder, which was accessible only to authorized members of the research team. These coded responses were then used for the analysis phase, thus decoupling any data from identifiable individuals and guaranteeing that no person or entity outside the research team could link the responses to a participant's identity.

Even though we cited certain comments from faculty members in the research, the utmost care was taken to ensure that these quotes could not be traced back to any individual. Furthermore, the survey was designed such that any partial responses could not be tied back to an individual participant due to the lack of identifiable information. This meant survey participants could withdraw from the study at any point without their previous participation being recognized.

B. Data collection

For this investigation, data were gathered through two sources. The first was in-depth, one-on-one, semistructured interviews with eight faculty members from the Physics Department at Auburn University. The interviewer was a tenure-track faculty member (E. B.) to ensure faculty felt comfortable expressing their views. All interview

TABLE II. Summary of data-collection instruments, participants, and methods from Auburn University's Physics Department.

Data source	Details
Interview participants	Eight faculty members from the Physics Department at Auburn University
Survey respondents	23 faculty members (96% participation within the department)
Communication medium for invitation	Email to the faculty mailing list
Interviewee composition	All were tenure-track faculty from a range of subdisciplines and career levels
Survey composition	21 tenure-track faculty members, 2 research professors
Faculty specialization	6 theorists, 17 experimentalists
Faculty with GDE-like exam experience	15 faculty members

participants were tenure-track faculty from a range of subdisciplines and career levels. Subsequently, a survey based on the interview findings was conducted, with 23 of 24 tenure-track faculty members responding to the survey. All members of the physics faculty were invited to participate in the interviews and survey through email communication sent to the faculty mailing list. Among the faculty members surveyed, there were 21 tenure-track faculty members and 2 research professors, 6 of the faculty members were theorists, while 17 were experimentalists. It is worth noting that among the surveyed faculty members, 15 had taken a qualifying exam similar to GDE during their graduate school education. These descriptions are summarized in Table II.

The interviews were semistructured and designed to offer a forum for faculty to share their experiences and insights regarding the GDE. Seven key themes were explored: faculty experiences with the GDE, perceptions of the GDE's purpose, the faculty role in student preparation for the GDE, thoughts on the consequences for students who fail the GDE, faculty members' personal experiences taking the GDE or an equivalent examination as students, their vision for the role of the GDE within an ideal graduate program, and any other GDE-related aspects that had not been addressed in the interview. The questions on the follow-up survey included themes related to personal experiences with the GDE, views on the grading methodology, perceptions of difficulty levels, and opinions about the GDE's necessity and purpose. This approach of utilizing both interviews and surveys allowed for a comprehensive gathering of data, integrating the qualitative depth of the interview responses with the breadth of the survey responses. Table III provides a detailed overview of the thematic areas explored within both the interview and survey protocols, as well as specific questions that correspond to those themes.

C. Data analysis

Our data analysis strategy deviated from traditional grounded theory methods [31], such as line-by-line coding, and instead adopted a comprehensive coding approach [32,33]. Line-by-line coding may lead to overlooking broader thematic insights. Our strategy was to capture the overarching message by interpreting each interview in its entirety. This comprehensive approach allowed us to identify the primary themes and ideas emerging from the interviews [34]. Although this method may have risked missing out on more detailed elements that line-by-line coding might have unveiled, our aim was more attuned to the larger thematic landscape rather than intricate details [35,36].

Grounded theory coding generally comprises three stages: open coding, axial coding, and selective coding. Open coding represents the initial phase where raw data are broken down into discrete parts, closely examined, and compared for similarities and differences. In this case, our coding was initially broken down by individual questions in the interview or on the survey. The axial coding phase involves reassembling the data fractured during open coding by making connections between categories. The final stage, selective coding, involves integrating and refining the theory. The central category is selected and systematically related to other categories. Validation procedures are used to ensure the theoretical framework fits and is relevant to the data [36]. In the selective coding phase, we synthesized these key themes into an integrative theory that represented faculty perceptions and experiences with the GDE.

To ensure the reliability and validity of our coding process, a reliability check was conducted by E. B. Following the initial coding and theme identification, E. B. independently reviewed the coded data to assess the consistency and accuracy of the codes and themes identified by S. B. Any disagreements or uncertainties were discussed and resolved through collaborative discussions between S. B. and E. B. We used a constant comparative method [24] to analyze the survey and interview data together. This iterative process of cross-examining data helped refine our research questions and allowed us to develop a theory that describes faculty perceptions and experiences about the GDE at Auburn University.

TABLE III. Overview of detailed exploration of thematic areas addressed in interviews and survey protocols on GDE.

Research question	Theme	Interview questions	Survey questions
RQ1: What are faculty members' perceptions regarding the purpose and necessity of written qualifying exams?	Understanding of the GDE's purpose	What do you think is the purpose of the GDE?	In two sentences or less, describe what you think the students think the purpose of the GDE is.
		Do you think this is the same view the students have? Why or why not?	Do you think this is the same view the students have? Why or why not? In your mind, what is the purpose of the GDE? Select all that apply. ☐ To assess students' knowledge in the fundamental areas of physics. ☐ To determine whether or not a student will be successful in the Ph.D. program. ☐ To take the burden of kicking students out of the program off of individual (or a smaller number of) faculty ☐ For the students to learn and grow as a cohort through shared experience. ☐ To validate the students and help them recognize their progress from undergraduates to graduate students. ☐ To cover gaps in material covered by the core courses. ☐ As an additional check to compensate for uncertainty in the admissions process. ☐ It is just a barrier to students starting research. ☐ Other (please describe).
RQ2: How do faculty members perceive their role in preparing students for the GDE, and what are their views on the relationship between such qualifying exams and the coursework in a graduate program?	Faculty responsibilities for GDE preparation	Do you think it is necessary to take the GDE after passing the required core courses? Why or why not	In two sentences or less, describe what you think the role of faculty is in preparing students for the GDE.
		Do you think the quality of core courses is consistent each year? What do you think the role of faculty in helping students prepare for the GDE is?	I think the quality of the core courses is consistent from year to year. ∘ Disagree

(Table continued)

TABLE III. *(Continued)*

Research question	Theme	Interview questions	Survey questions
		Do you think if the department created a GDE prep course that it would be helpful for students?	∘ Natural ∘ Agree I think the core courses should prepare the students for the GDE. ∘ Disagree ∘ Neutral ∘ Agree I think a summer GDE preparatory course would be beneficial for the students. ∘ Disagree ∘ Neutral ∘ Agree I think the GDE is necessary even if a student has been successful in the core courses. ∘ Disagree ∘ Natural ∘ Agree
	Grading system of GDE	On average, how many questions do you (faculty) provide in the development of the GDE each year? And if none then why?	How many GDE questions do you write each semester? ∘ 0 ∘ 1 ∘ 2+ How do you grade the GDE (for example, do you create a rubric? What do you give the most points for?). Do you think there is a difference in the difficulty of the fall and spring GDE? ∘ No ∘ It depends ∘ Yes Do you think the spring GDE is graded more leniently than the fall GDE? ∘ No

(Table continued)

TABLE III. *(Continued)*

Research question	Theme	Interview questions	Survey questions
			∘ It depends
			∘ Yes
RQ3: How do faculty members perceive the effectiveness of qualifying exams in measuring students' knowledge, skills, and potential success in physics?	Personal experience with GDE as faculty students and as a student	Tell me about your experience with the GDE as a faculty member. On average, how many questions do you (faculty) provide in the development of the GDE each year?	I took a qualifying exam similar to the GDE in graduate school. (Closed-ended question).
		Do you think the GDE is a good measure of how successful a student will be as a graduate researcher? Why or why not?	∘ No
		Could you describe your experience taking the equivalent of the GDE when you were a graduate student?	∘ Yes
		How exactly did you prepare for your GDE?	The process of studying for and taking that exam made me a better physicist.
		How did you feel during the preparation period?	∘ No
		Do you feel like the GDE was a fair evaluation of your physics knowledge/skills? Why or why not?	∘ Yes
		Do you feel like the experience of taking the GDE made you a better physicist?	Do you think the GDE is an indicator of how successful a student will be as a researcher?
		Other than providing you an opportunity to complete your graduate program, how much has the GDE contributed to your success in your physics career?	∘ No
			∘ Yes
	Viewpoints on student failure in GDE	What do you think should happen to students who do not pass the GDE?	The next few questions are about what happens if a student fails the GDE.
		Would your perception of a student as a physicist change If they did not pass the GDE? If so, how?	If a student did not pass the spring GDE, I would recommend they continue their graduate studies in another **physics** program.
		Would you recommend them to continue with their doctoral studies in another program?	∘ Disagree
		What advice would you provide to these students on their academic careers?	∘ Neutral
			∘ Agree

(Table continued)

TABLE III. *(Continued)*

Research question	Theme	Interview questions	Survey questions
			If a student did not pass the spring GDE, I would recommend they continue heir graduate studies in a **different field**. ∘ Disagree ∘ Neutral ∘ Agree
			If a student did not pass the spring GDE, I would recommend they seek a position in **industry**. ∘ Disagree ∘ Neutral ∘ Agree
			I have lost a research student because they did not pass the GDE ∘ No ∘ Yes
			I think this person still would have been successful as a researcher. ∘ No ∘ Unsure ∘ Yes
RQ4: What do faculty members suggest as modifications to or alternatives for qualifying exams?	Ideal vision of the GDE	If you had complete control over the design of the graduate program, describe the role (if any) that the GDE would play in that program?	What is your general view of the GDE in its current form?
		If you kept the GDE, what would you change about the content or format of the exam?	∘ Favorable
		If you would abolish the GDE, please indicate why. What (if anything) would you replace it with?	∘ Neutral
		Do you think the GDE is a fair evaluation of a student's physics knowledge/skills? Why or why not?	∘ Unfavorable
			Below, please write what changes you would make to the current GDE process, if any.

(Table continued)

TABLE III. *(Continued)*

Research question	Theme	Interview questions	Survey questions
			These are some changes to the GDE that have been suggested by faculty members and students. Please select all that you think would be acceptable. ☐ A comprehensive multiple choice exam testing basic understanding of fundamental physics ideas. ☐ Allowing the students to take the GDE every year until they pass all four exams. ☐ Put more emphasis on the Oral Exam and make the format and requirements for that more strict. ☐ Having an oral exam focused on knowledge and skills in the students' chosen research area (separate from their thesis proposal). ☐ Doing a more holistic evaluation of each student including performance in courses and on the GDE, research progress, as well as the student's personal circumstances. ☐ Allowing students who receive an A in the core course not to take the GDE in that subject. ☐ Preparing common final exams for the core courses, similar to what is done for the service courses. ☐ None of these are acceptable, but I still suggest some change (describe below). ☐ No modifications are necessary or wise.
Additional comments or insights		Is there anything else you would like to talk to us about the GDE that you don't feel we have covered here?	What are the potential negative effects of not having a GDE? What are the potential positive effects of not having a GDE?

III. RESULTS

The study examined faculty members' perceptions regarding the graduate doctoral examination (GDE) after completion of core courses in a physics graduate program. The findings revealed skepticism toward the current format of the GDE as a fair evaluation method, with a plurality of faculty members expressing an unfavorable opinion. Faculty members highlighted concerns about the GDE's ability to measure students' physics knowledge and skills accurately. Many faculty members considered the assessment of students' knowledge in fundamental areas of physics as the primary purpose of the GDE. However, faculty members also raised concerns about the quality and consistency of core courses, emphasizing the need for alignment with the GDE. While some faculty members believed the GDE could predict a student's success in the Ph.D. program, the majority did not view it as an accurate indicator. Most participants expressed a desire to retain the GDE in some form but suggested modifications to improve its effectiveness.

A. RQ1: What are faculty members' perceptions regarding the purpose and necessity of written qualifying exams?

Faculty members articulated several different purposes that the GDE serves. In the interviews, the responses demonstrated a focus on student evaluation and personal development, with most indicating the primary purpose was to assess the students' knowledge in physics and others identifying personal progress and distinction between graduate and undergraduate as essential goals. Furthermore, a few faculty considered the GDE's purpose to be confirming students' eligibility for a degree in physics, while others viewed it as a mechanism to off-load the task of dismissing students from individual faculty's shoulders. One faculty member saw the GDE as a tool to cover the gaps in knowledge or to compensate for uncertainty in the admissions process. These diverse perspectives on the GDE's purpose collected during interviews formed the basis for our subsequent faculty-wide survey.

The survey provided a broader view of the faculty's opinions (see Fig. 1). The most popular survey response, selected by 18 faculty, mirrored the interview data that highlighted the GDE's primary purpose as assessing students' knowledge in fundamental areas of physics. Thirteen of the survey respondents affirmed the GDE's role in distributing the faculty burden of removing students from the program. This finding implies that a considerable portion of the faculty sees the GDE as a procedural tool to manage student progress and academic continuity. Fifteen respondents considered the GDE an additional check to compensate for uncertainty in the admissions process. Ten faculty believed it helped validate students' transition from undergraduate to graduate students, signifying its role in promoting students' academic and personal growth. Interestingly, only three faculty saw the GDE as a tool to cover gaps in material covered by the core courses, and only four saw it purely as a barrier to students starting research. Finally, nine respondents agreed that one of the purposes of the GDE was to facilitate collective student growth through shared experiences.

While there are multiple viewpoints on the GDE's purpose, the faculty appear to agree on its fundamental role in assessing and fostering students' understanding and growth in the field of physics. The GDE, as seen by the faculty, serves as a useful tool in ensuring academic rigor, managing student progress, and smoothing transitions in the physics department. However, it was found that while 14% believed the GDE could predict a student's success in the Ph.D. program, 66% did not view it as an accurate indicator of future success as a researcher. Curiously, this ratio was the same for both theorists and experimentalists. We note that we did not specify a fixed definition of "success," as that tends to vary widely by field and even by individual. For example, one of the interview subjects considers two published papers or more as a "successful"

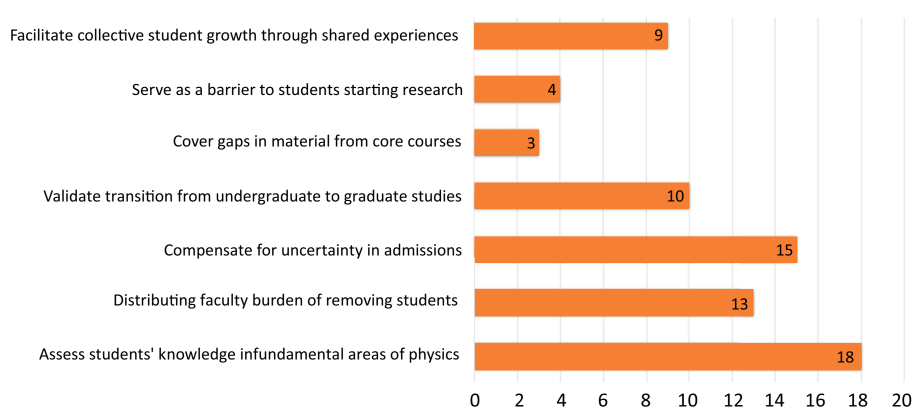

FIG. 1. Survey results from 23 faculty members. Each faculty member could select more than one choice from the survey options regarding purpose. The number of faculty members who select each option is shown in the bar plot.

Ph.D. student, while E. B. has different expectations for publication, etc., depending on each student's career goals.

B. RQ2: How do faculty members perceive their role in preparing students for the GDE, and what are their views on the relationship between such qualifying exams and the coursework in a graduate program?

Faculty members emphasized the need for careful evaluation of coursework and alignment with the GDE, as they believe the core courses should adequately prepare students for the examination. When asked in the interview whether they thought the quality of core courses was consistent each year, all eight faculty members interviewed expressed their belief that the core courses were not consistent year to year. Furthermore, a majority of interviewees indicated a belief in the importance of the GDE as an essential follow-up to the core courses. However, some disagreed, suggesting that passing the required core courses might be a sufficient indication of mastery of the material.

Individual comments from faculty members added nuance to these perspectives. One interviewee suggested a reimagined version of the GDE: "Yes, I think we need to have something like GDE. As it stands, maybe not. Maybe something like an admission test." Another faculty member saw the purpose of core courses as a stepping stone to the GDE: "Yes, because the core courses' purpose is trying to get them into a graduate-level kind of mentality and being able to do work at that level." Other respondents supported the GDE in a different form, focusing on problem solving and mastering core concepts: "Yeah, I think so in the sense again it doesn't have to be in the form of a test, to improve problem-solving and approach new problems and core fundamental concepts." One faculty member suggested that core course assessments should be leveraged to evaluate students' qualifications but acknowledged the inherent gray areas: "Testing in core courses should be used to evaluate students' academic credentials to continue but acknowledges that it will never be black and white." The final interviewee emphasized the unique role of the GDE: "Yes, the GDE is different from core courses because it provides a set of paths to follow without a guide." The survey further substantiates these findings. When asked to evaluate the statement, "As a faculty member, I think the quality of the core courses is consistent from year to year," the majority of the respondents expressed uncertainty. More specifically, 12 faculty members (52%) were unsure, while 8 faculty members (35%) disagreed with the statement. Only three faculty members (13%) agreed.

Survey respondents were also asked to agree or disagree with the statement, "I think the core courses should prepare the students for the GDE." Around 65% of faculty agreed with this statement and 26% of faculty were ambivalent or uncertain about this argument, while only 2 faculty members (9%) disagreed, stating that core courses and GDE preparation need not be intertwined. In addition, when asked about the faculty's role in preparing students for the GDE, 10 faculty members mentioned coaching students in problem-solving skills, deficiencies, or study skills (see Fig. 2). Eleven faculty members believed aligning course material with the GDE was part of their responsibility. Two faculty members thought that explaining the reasoning behind the GDE was part of their role, while a single faculty member mentioned the idea of lowering the bar for passing. In summary, while a small fraction of the faculty perceives the quality of core courses to be consistent each year, the majority expresses either disagreement or uncertainty. Additionally, the faculty do seem to agree that they play a role in preparing the students for the process.

C. RQ3: How do faculty members perceive the effectiveness of qualifying exams in measuring students' knowledge, skills, and potential success in physics?

During the interviews, faculty members provided valuable insights into their experiences and the perceived impact of the GDE. One faculty member highlighted the concentrated effort and mild apprehension associated with the exam, stating, "I think the concentrated effort of just, you know, sitting down, being mildly afraid of the whole process," as a justification for how the exam facilitated their

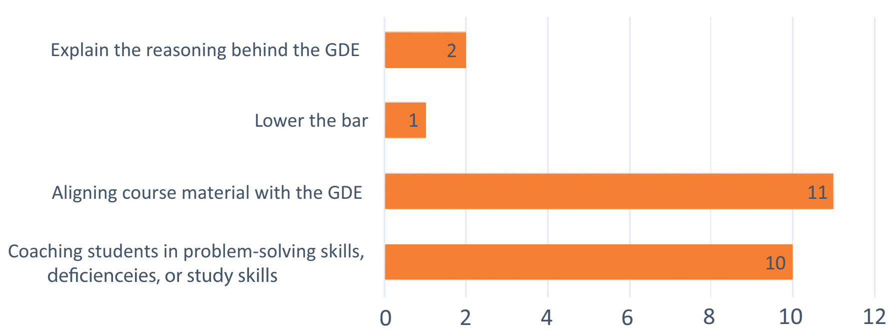

FIG. 2. Survey results of 23 faculty members. Each faculty member could select more than one choice from the survey options regarding the role of faculty in preparing students for GDE. The number of faculty members who select each option is shown in the bar plot.

personal growth as a physicist. This faculty member also suggested that the environmental stress induced by the exam actually made them more productive. They further emphasized the positive aspects of having a "forcing function" that drives one to work harder and accomplish more. Another faculty member expressed how the process of studying for and taking the exam had made them a better physicist. They emphasized the importance of perseverance, stating, "sitting down and forcing myself to follow through on each and every problem to the end….I developed both an appreciation and a protocol." However, it is important to note that not all faculty members attributed their growth solely to the GDE. One faculty member highlighted the significance of research experiences, problem solving in the lab, and collaborating with different investigators, stating, "what really made me a better physicist are the experiences and research that I have, and solving problems, real problems, in the lab, and working with different investigators, and seeing how they solved problems."

While some faculty members recognized the GDE as a valuable experience that fostered personal and intellectual development, others emphasized the importance of broader research experiences in shaping their skills and understanding of physics. The survey revealed that among the faculty members who had experienced a similar format of the GDE in their Ph.D. program, 83% expressed that it had positively contributed to their development as physicists.

Contrasting with the prevailing viewpoint on the importance of the GDE as an evaluative tool, we explored its predictive value for future research success. We asked faculty members, both in interviews and surveys, about the correlation between performance on the GDE and subsequent research achievements, while considering other potential influences such as prior research experience. All interview responses asserted that passing the current GDE was not indicative of a student's future success as a researcher. This stance was also reflected in the subsequent survey, where a majority of 63% disagreed with the notion that GDE performance could accurately predict research success while 37% agreed. Nearly 40% of faculty members (9 out of 24) reported losing a research student due to failing the GDE, though only 3 of them were certain the students would have been unsuccessful if the student were allowed to stay in the program.

The faculty's attitudes toward student outcomes following an unsuccessful GDE attempt were also explored in the survey. When presented with the statement, "If a student did not pass the spring GDE, I would recommend they continue their graduate studies in another physics program," a majority of faculty members (59%) expressed uncertainty. The remaining responses were divided between disagreement and agreement, with 9% expressing opposition, and 32% supporting the idea. When the scenario was adjusted to whether a student should continue their graduate studies in a different field after not passing the GDE, neutrality was again the most common response. Approximately, 68% of faculty were neutral, while disagreement and agreement were less common, accounting for 18% and 14%, respectively. Finally, regarding the statement, "If a student did not pass the spring GDE, I would recommend they seek a position in industry," a high degree of neutrality was also observed. Around 73% expressed a neutral stance, with disagreement and agreement both holding minor positions, each shared by 14% of faculty. These findings underscore a prevailing uncertainty within the faculty regarding the suggested paths for students who do not succeed in the GDE.

D. RQ4: What do faculty members suggest as modifications to or alternatives for qualifying exams?

The survey responses revealed that 26% of the faculty members expressed a favorable view of the GDE's current format, while 47% held an unfavorable opinion, and 26% remained unsure or undecided. The interview data further supported these findings, with most interviewees expressing the belief that the current format of the GDE is not a fair evaluation method. One faculty member emphasized the disparity in what the GDE measures, stating, "No, I don't think [it's fair]. I think it's evaluating their ability to retain information but not the concept." Another respondent agreed on the limitations of a singular point of evaluation: "No, a single point evaluation never is fair." A third respondent said: "No, the current format is a much better measurement of the math skills than actually the physics skills." Nonetheless, some respondents suggested that the GDE could potentially be a fair assessment, contingent on its execution: "No, [the] GDE can be a fair evaluation of a student's physics knowledge and skills, but it depends on how the test is executed."

When asked in the interview what changes they might make, the faculty gave varied responses (see Fig. 3). Some faculty members advocated for the complete elimination of the GDE process while others proposed modifications to its structure or grading process. Many of the faculty expressed

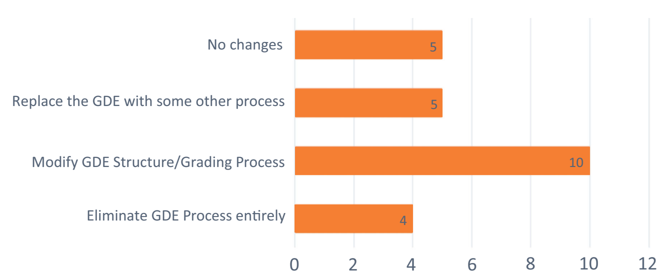

FIG. 3. Survey results of 23 faculty members. Each faculty member could select more than one choice from the survey options regarding changes to GDE. The number of faculty members who select each option is shown in the bar plot.

TABLE IV. Overview of survey analysis of faculty members' suggestions for GDE modifications.

Proposed change	Percentage of total responses	Number of votes
Conduct an oral examination on students' research area	20%	12
Implement a more holistic student evaluation	18%	11
Grant exemption for students with A in corresponding core	13.33%	8
Prepare common final exams for core courses	13.33%	8
Strengthen emphasis on the Oral Exam	11.67%	7
Transform GDE into a comprehensive multiple-choice exam	6.67%	4
Allow students to take the GDE annually until passing all four exams	6.67%	4

a desire to replace the GDE with a different process altogether, signaling a potential need for a broader reimagining of the examination. Only two of the faculty expressed satisfaction with the current format, advocating for no changes. One faculty member stated, "The exam should be more project-oriented rather than focused on memorizing equations." Another faculty member suggested implementing a stricter oral defense process, where graduate students would work with a committee for a year and their success would be evaluated. As expressed by the faculty member, "A stricter oral defense, where a graduate student works with a committee for a year and then evaluates their success. This would give more weight to courses and potentially require restructuring of exams." Furthermore, another faculty member recommended "having students present their research progress at the end of the first or second year to demonstrate their dedication to the program."

When questioned about potential modifications to the process in the broader survey, results displayed some shifts. Support for the elimination of the GDE process was at 17% (see Table IV and Fig. 4), while the call for modifying the GDE structure or grading process was larger at 41%. This suggests growing faculty interest in refining the existing process rather than removing it. Around 21% suggested a replacement for the GDE, and 21% favored no change.

In the survey, the majority of faculty expressed a desire to retain the GDE but suggested modifications to improve its effectiveness. Only three individuals opposed any alterations to the current system. The faculty put forth several potential alterations to improve the GDE. The most widely supported reform, endorsed by 12 respondents, was to conduct an oral examination centered on knowledge and skills in the student's chosen research area, separate from their thesis proposal. Another suggestion, supported by 11 respondents, was a call for a more holistic student evaluation that considers performance in courses, on the GDE as it currently is administered, research progress, and the student's personal circumstances. Eight faculty members supported granting an exemption from the GDE for students who score an A in the corresponding core course. An equal number recommended the preparation of common final exams for the core courses, mirroring the approach taken for service courses in this department. Seven respondents advocated for a stronger emphasis on the oral exam taken later in the program (but separate from the defense), suggesting stricter formatting and requirements. Four faculty supported transforming the GDE into a comprehensive multiple-choice exam to test fundamental physics concepts. The same number of participants also supported allowing students to take the GDE annually until they pass all four examinations.

While the GDE maintains some level of support among faculty members, there is a significant consensus toward its reform. These findings demonstrate the

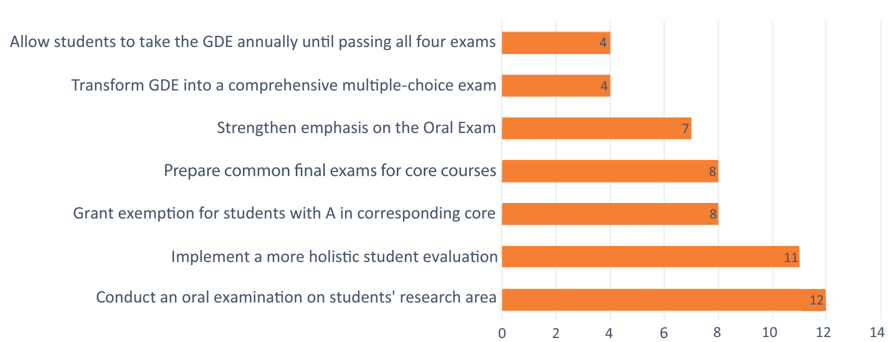

FIG. 4. Survey results of 23 faculty members. Each faculty member could select more than one choice from the survey options regarding more suggestions to change current format of GDE. The number of faculty members who select each option is shown in the bar plot.

faculty's diverse perspectives on potential changes to the GDE, underscoring the need for a more in-depth exploration of the strengths and weaknesses of the current process.

IV. DISCUSSION

Our exploration into faculty members' perceived purpose of GDE illuminated a variety of viewpoints. Despite the diversity, many faculty agreed on the GDE's fundamental role in assessing students' knowledge and their preparedness for a degree in physics and fostering academic growth. Moreover, many faculty members perceive the implementation of the GDE as helpful in validating the transition from undergraduate to graduate studies. The findings align with Estrem and Lucas which indicated that qualifying exams are academic mechanisms designed to ensure Ph.D. candidates' competency and preparedness before embarking on research [7]. It is also consistent with Furstenberg and Nichols-Casebolt and Kostohryz, who point out the role of the exam as a quality control measure, targeting students' understanding of key theories, topics, and methodologies in the discipline [8,9].

Notably, based on our survey and interview results, the GDE was seen as an essential tool to facilitate difficult decisions about student continuation in the program. Schmidt *et al.* also reported that these evaluations represent a blend of tradition and standard setting, simultaneously acting as a badge of "rigorous" academic training and a filtering mechanism for doctoral candidates [10]. Interestingly, while most faculty members believe the GDE does not accurately depict a student's research capabilities; they still consider it an indicator of "readiness" for research. This suggests that while the GDE exam is a crucial stepping stone, it does not encompass all the necessary conditions for research proficiency. This is also supported by the responses that suggest some faculty members are not convinced that students who fail the GDE would have been unsuccessful in research.

The faculty responses imply that fundamental "knowledge" of physics is necessary but not sufficient for research success. This finding may not be surprising given that the respondents were mostly experimentalists, and that the "skills" and "knowledge" used in experimental physics are rarely addressed in core graduate coursework [37]. Still, it is curious that demonstrating "knowledge" of physics by being able to successfully solve textbook-style questions in a timed setting would somehow be a prerequisite for learning how to troubleshoot equipment, for example. An alternative explanation could be that faculty expect students to be fluent in both settings, even if they are not directly related. For example, if a student will go on to be a faculty member, it might be expected that they are able to "do physics" the way that it is done in the classroom *in addition* to being proficient in research. Further studies about instructor epistemology related to ways of knowing in physics graduate programs would provide useful information about what it means to be a physicist.

Many faculty still agreed that the GDE would be necessary after the core coursework. However, significant doubts arose among faculty members regarding the consistency of core courses and their alignment with the GDE. Many faculty members believed that the core courses' consistency and GDE should be more aligned. This dissonance is mirrored in the survey data. Overall, the results indicate that, while most faculty believe that some form of summative assessment for students is necessary, there is substantial misalignment between assessment methods, teaching methods, and learning outcomes for the graduate program. Faculty concerns about varying quality in core courses each year point to potential reliability problems. If these courses aim to equip students for the GDE, any inconsistency might undermine its efficacy. The mixed opinions on the need for the GDE after core courses also raise doubts about these courses' effectiveness in preparing students. This inconsistency conflicts with the principles of backward design, which advocates setting an end goal, such as the GDE, and subsequently tailoring courses to achieve that objective [38–41].

The experiences and training the faculty members undergo during their graduate studies play a pivotal role in forming their beliefs about academic assessment [42,43]. This might explain why faculty members have specific views on the graduate Ph.D. examination and its alternatives, as they draw from their foundational experiences during their doctoral programs. This suggests that departmental cultural norms surrounding the GDE are largely shaped by the individual experiences of faculty members with such examinations. The precedence of these exams in physics doctoral education may explain the reason most faculty view this as a necessary tool. However, the wide variety of ways in which such exams are administered and evaluated might explain the diversity of faculty opinions on how such exams should be executed.

A notable finding is the relative lack of agreement between faculty on potential alternatives to this process. Approximately, half the faculty seem to think that a more holistic evaluation method that encompasses other aspects of students' performance, such as research, would be a better method. They do, however, still believe that some kind of summative evaluation is needed to ensure the students are prepared for the rest of their degree. This makes the likelihood of second-order change in the program somewhat unlikely. Our findings align with observations about the complexities and challenges inherent in starting change within STEM disciplines [44]. Discussion about the nonlinearity of the change process in STEM education by Henderson *et al.* is visible in the division observed among faculties regarding potential modifications to the GDE.

V. LIMITATIONS AND FUTURE WORK

Our study's primary focus has been the faculty, rendering it inherently biased toward this group's perceptions. Not only are these responses likely subject to survival bias [45] but also omit the student perspective. Students, being the group most impacted by the GDE process, will have unique insights that will enrich our understanding. We acknowledge this gap and are working to bridge it; a dedicated analysis centered on the student perspective is in progress and will be discussed in a forthcoming publication.

Another limitation is that we only collected information from a single department. While this department's dynamics and challenges offer valuable insights, they may not be wholly reflective of the diverse range of physics faculties dispersed throughout the United States. The inherent cultural, administrative, and pedagogical variations across institutions can lead to divergent experiences and practices related to the GDE. Hence, our study's findings should be interpreted with caution when considering their applicability to other departments or institutions. Despite these limitations, we were able to achieve a nearly 100% response rate within an individual department, which shows the wide variety of viewpoints that can exist in an environment that might be considered homogeneous in a larger-scale survey investigation. Indeed, there were several points on which the interviewed faculty were in complete or near-complete agreement, but the broader survey demonstrated more hedging or disagreement. This could be due to social desirability bias in the interview setting or could reflect an issue with selection bias that would certainly affect larger-scale data collection efforts [46]. For example, a national survey might reflect a more progressive view on qualifying exams due to response bias by those most motivated to change these processes.

This study represents the first step in a longer-term case study of departmental change. For example, in response to the results of the interviews and survey, the department has assembled two committees. The first committee consists of the people responsible for writing the GDE and the core course instructors, who will oversee writing comprehensive learning objectives for both the exam and the courses. These objectives will be distributed to all students and faculty to assist faculty in the process of writing exam questions, and the students in studying. The second committee oversees codifying (i) the purpose of the GDE and (ii) the form it might take in the future. The establishment of these two committees in the department with different focuses indicates the department's initially conservative approach to navigating the challenges and tension associated with the GDE.

While the faculty are grappling with the issue of the GDE, the graduate students in this department have become more organized and vocal about their opinions of departmental practices. For example, the graduate student association was completely restructured following this data collection, and graduate student representatives are now privy to a wider range of decision making in the department. Indeed, this preliminary work was also a driving factor in this university joining a nationwide effort to promote inclusion in physics graduate programs [47]. The GDE has been at the center of many of these discussions that involve a variety of departmental stakeholders including leadership, faculty, staff, and students. This work is also forming the basis for a longer-term project that investigates how the GDE may drive students out of the program and the effects it has on them.

VI. CONCLUSION

Faculty perspectives on the GDE in the physics department at Auburn University reveal a complex tapestry of opinions. There is consensus on the GDE's role in gauging students' content knowledge, but its ability to predict future research success remains contested. Many faculty members argue that while the GDE maintains academic standards, it may not reliably forecast a student's expertise in research. The interplay between core courses and the GDE is also a topic of debate, with concerns about the year-to-year consistency of core courses and their alignment with the GDE. Moreover, suggestions point toward a more holistic approach to the GDE, including project-oriented assessments and an emphasis on research presentations.

The split in faculty views on the GDE's reform underscores the challenges of instigating meaningful change. Drawing parallels with STEM discipline transformation literature, navigating this intricate landscape requires a collaborative strategy, harmonizing the myriad of faculty opinions. The ultimate aim is to recalibrate the GDE, ensuring it aligns with broader educational objectives and provides a robust, equitable evaluation of students' readiness in physics, without diminishing its academic rigor.

[1] *Envisioning the future of doctoral education: Preparing stewards of the discipline*, edited by C. M. Golde and G. E. Walker, Carnegie essays on the doctorate (Jossey-Bass, San Francisco, CA, 2006).

[2] R. Liera, A. J. Rodgers, L. N. Irwin, and J. R. Posselt, Rethinking doctoral qualifying exams and candidacy in the physical sciences: Learning toward scientific legitimacy, Phys. Rev. Phys. Educ. Res. **19**, 020110 (2023).

[3] S. K. Gardner, Student and faculty attributions of attrition in high and low-completing doctoral programs in the United States, Higher Educ. **58**, 97 (2009).

[4] P. Green and J. Bowden, Completion mindsets and contexts in doctoral supervision, Qual. Assur. Educ. **20**, 66 (2012).

[5] B. E. Lovitts, *Leaving the Ivory Tower: The Causes and Consequences of Departure from Doctoral Study* (Rowman & Littlefield Publishers, Lanham, MD, 2001).

[6] J. D. Nyquist and D. H. Wulff, *Working Effectively with Graduate Assistants* (Sage Publications, Inc., Thousand Oaks, 1996).

[7] 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).

[8] A. L. Furstenberg and A. Nichols-Casebolt, Hurdle or building block: Comprehensive examinations in social work doctoral education, J. Teach. Soc. Work **21**, 19 (2001).

[9] K. Kostohryz, The doctoral comprehensive examination in counselor education: Faculty members' perception of its purposes, J. Counselor Prep. Superv. **8**, 1 (2016).

[10] 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).

[11] R. A. Beebe, W. A. Noyes, B. Riegel, W. G. Young, S. C. Lind, and E. M. Billings, Philosophy of graduate training at Ph. D. Level, Chem. Eng. News **26**, 165 (1948).

[12] N. Ponder, S. E. Beatty, and W. Foxx, Doctoral comprehensive exams in marketing: Current practices and emerging perspectives, J. Market. Educ. **26**, 226 (2004).

[13] S. Gelmez Burakgai and A. Yildirim, Journey into doctoral candidacy: A grounded theory of doctoral qualification exam process, Croatian J. Educ. **19**, 63 (2017).

[14] American Center for Physics, Graduate education in physics: Which way forward?, in *Proceedings of a Conference to Discuss the Status and Future of Graduate Education in Physics* (American Physical Society and American Association of Physics Teachers One Physics Life, College Park, MD, 2008).

[15] 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).

[16] 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).

[17] 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. Stud. Math. **50**, 177 (2002).

[18] A. J. MacLachlan, Preservation of Educational Inequality in Doctoral Education: Tacit Knowledge, Implicit Bias and University Faculty. Research & Occasional Paper Series: CSHE. 1.17, Center for Studies in Higher Education (2017).

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

[20] *Power and Privilege in the Learning Sciences: Critical and Sociocultural Theories of Learning*, edited by I. Esmonde and A. N. Booker (Taylor & Francis, New York, NY, 2016).

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

[22] S. Gaukroger, *Objectivity: A Very Short Introduction* (Oxford University Press, New York, 2012), Vol. 316.

[23] D. L. Reinholz, R. L. Matz, R. Cole, and N. Apkarian, STEM is not a monolith: A preliminary analysis of variations in STEM disciplinary cultures and implications for change, CBE Life Sci. Educ. **18**, mr4 (2019).

[24] Y. Chun Tie, M. Birks, and K. Francis, Grounded theory research: A design framework for novice researchers, SAGE Open Med. **7**, 2050312118822927 (2019).

[25] R. Thornberg, Informed grounded theory, Scand. J. Educ. Res. **56**, 243 (2012).

[26] M. Palmer, M. Larkin, R. O. de Visser, and G. Fadden, Developing an interpretative phenomenological approach to focus group data, Qual. Res. Psychol. **7**, 99 (2010).

[27] A. Alase, The interpretative phenomenological analysis (IPA): A guide to a good qualitative research approach. Int. J. Educ. Literacy Stud. **5**, 9 (2017).

[28] K. Reid, P. Flowers, and M. Larkin, Exploring lived experience, Psychologist **18**, 20 (2005), https://www.bps.org.uk/psychologist/exploring-lived-experience.

[29] A. E. Austin, Faculty cultures, faculty values, New Directions Institut. Res. **1990**, 61 (1990).

[30] R. Donnelly, Perceived impact of peer observation of teaching in higher education, Int. J. Teach. Learn. Higher Educ. **19**, 117 (2007), https://eric.ed.gov/?id=EJ901290.

[31] K. Charmaz, *Constructing Grounded Theory*, 2nd ed. (Sage, Thousand Oaks, CA, 2014).

[32] M. Williams and T. Moser, The art of coding and thematic exploration in qualitative research, Int. Manag. Rev. **15**, 45 (2019).

[33] A. Strauss and J. Corbin, *Basics of Qualitative Research: Grounded Theory Procedures and Techniques* (Sage Publications, Thousand Oaks, CA, 1990).

[34] J. Saldana, *The Coding Manual for Qualitative Researchers* (Sage, Thousand Oaks, CA, 2015).

[35] R. Thornberg and K. Charmaz, Grounded theory and theoretical coding, in *The SAGE Handbook of Qualitative Data Analysis* (Sage Publications, Thousand Oaks, CA, 2014), Vol. 5, pp. 153–169.

[36] V. Braun and V. Clarke, Reflecting on reflexive thematic analysis, Qual. Res. Sport Exercise Health **11**, 589 (2019).

[37] C. Wieman, Comparative cognitive task analyses of experimental science and instructional laboratory courses, Phys. Teach. **53**, 349 (2015).

[38] D. J. Rybacki and D. L. Lattimore, Assessment of undergraduate and graduate programs, Public Relat. Rev. **25**, 65 (1999).

[39] A. W. Astin, Student involvement: A developmental theory for higher education, in *College Student Development and Academic Life* (Routledge, London, 2014), pp. 251–262.

[40] F. Martin, A. Ritzhaupt, S. Kumar, and K. Budhrani, Award-winning faculty online teaching practices: Course design, assessment and evaluation, and facilitation, Internet Higher Educ. **42**, 34 (2019).

[41] *STEM project-based learning: An integrated science, technology, engineering, and mathematics (STEM) approach*, edited by R. M. Capraro, M. M. Capraro, and J. R. Morgan (Springer Science & Business Media, Rotterdam, Netherlands, 2013).

[42] A. E. Austin, Preparing the next generation of faculty: Graduate school as socialization to the academic career, J. Higher Educ. **73**, 94 (2002).

[43] C. Wieman, Expertise in University teaching and the implications for teaching effectiveness, Evaluation training, Daedalus (Boston) **148**, 47 (2019).

[44] C. Henderson, A. Beach, and N. Finkelstein, Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature, J. Res. Sci. Teach. **48**, 952 (2011).

[45] J. N. Carpenter and A. W. Lynch, Survivorship bias and attrition effects in measures of performance persistence, J. Financ. Econ. **54**, 337 (1999).

[46] I. Krumpal, Determinants of social desirability bias in sensitive surveys: A literature review, Qual. Quant. **47**, 2025 (2013).

[47] B. Goldberg, G. Cochran, C. Henderson, M. Wittmann, D. Sachmpazidi, S. Woods, and E. Su, Inclusive Graduate Programs: An AGEP Pilot in Physics and Astronomy, Bulletin of the American Physical Society (2024).