

Standards needed? An exploration of qualifying exams from a literature review and website analysis of university-wide policies

Exploration of
qualifying
exams

19

Jacqueline E. McLaughlin and Kathryn Morbitzer
Division of Practice Advancement and Clinical Education and Center for Innovative Pharmacy Education and Research, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Margaux Meilhac and Natalie Poupart
Center for Innovative Pharmacy Education and Research, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Rebekah L. Layton
Office of Graduate Education, Eshelman School of Pharmacy, The University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, North Carolina, USA, and

Michael B. Jarstfer
Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Received 16 November 2022
Revised 26 April 2023
Accepted 18 June 2023

Abstract

Purpose – While known by many names, qualifying exams function as gatekeepers to graduate student advancement to PhD candidacy, yet there has been little formal study on best qualifying exam practices particularly in biomedical and related STEM PhD programs. The purpose of this study is to examine the current state of qualifying exams through an examination of the literature and exploration of university-wide policies.

Design/methodology/approach – The authors conducted a literature review of studies on qualifying exams and completed an external evaluation of peer institutions' and internal institutional qualifying exam requirements to inform our discussion of qualifying exams practices in PhD training at a research-intensive US institutions.

© Jacqueline E. McLaughlin, Kathryn Morbitzer, Margaux Meilhac, Natalie Poupart, Rebekah L. Layton and Michael B. Jarstfer. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

This work was supported in part by National Institutes of Health General Medical Sciences – Science of Science Policy Approach to Analyzing and Innovating the Biomedical Research Enterprise (SCISIPBIO) Award (GM-19-011) – 1R01GM140282-01.



Studies in Graduate and
Postdoctoral Education
Vol. 15 No. 1, 2024
pp. 19-33
Emerald Publishing Limited
2398-4686
DOI [10.1108/SGPE-11-2022-0073](https://doi.org/10.1108/SGPE-11-2022-0073)

Findings – This study identified the need for more research on qualifying exams to establish evidence-based best practices. The authors found a wide variety of qualifying exam formats, with little evidence in support for specific formats. The authors also found little evidence that student expectations are made clear. The lack of evidence-based best practices coupled with insufficient clarity for students has a real potential to disadvantage PhD students, particularly first generation, underrepresented minority, international and/or other trainees who are not privileged or socialized to navigate training environments with vague landmarks such as the qualifying exams.

Originality/value – There are very few studies that evaluate qualifying exams in US doctoral education, particularly in STEM fields, and to the authors' knowledge, there has been no analysis of campus-wide policies on qualifying exams reported. The lack of evidence for best practices and the need for to evaluate the implementation and effectiveness of qualifying exams are discussed.

Keywords Graduate education, Qualifying exam, Preliminary exam, Candidacy, PhD training

Paper type General review

Introduction

The Doctor of Philosophy (PhD) is the highest academic degree for Science, Technology Engineering and Mathematics (STEM) fields and requires completion of an intense, mentored research project. The most common structure for US doctoral programs includes four elements: 1–2 years of core and elective coursework in the discipline; a qualifying exam (QE); in the form of a dissertation (Goldman and Massy, 2000; Hartnett and Katz, 1977; National Academies of Sciences, Engineering, and Medicine, 2018; Walker *et al.*, 2008). While the specifics of each element vary by discipline and institution, they generally exist in some form in most programs (Austin, 2002; Hartnett and Katz, 1977).

Increasingly, careers of young PhD scientists are diverging from their thesis work, leading to calls for ensuring that core competencies of PhD programs include translatable skills (e.g. leadership, communication, management) in addition to technical and discipline-specific competencies (e.g. scientific knowledge, technical skills, data interpretation) (McLaughlin *et al.*, 2019; Sinche *et al.*, 2017). As PhD training programs are not typically accredited by a professional organization in contrast to many professional degree programs, there is no common list of competencies or core outcomes defined. Instead, each program is tasked with defining their own core competencies and program outcomes. The 2018 report from the National Academies of Sciences, Engineering and Medicine Committee, *Revitalizing Graduate STEM Education for the 21st Century*, identified two overarching competencies: “Develop Scientific and Technological Literacy and Conduct Original Research” and “Develop Leadership, Communication, and Professional Competencies” (National Academies of Sciences, Engineering, and Medicine, 2018). Despite this attempt at defining common competencies, it remains unclear if or how these have been adopted by PhD programs and the extent to which PhD programs have defined their own core competencies – and corresponding assessments.

Our research team is pursuing a broad research program exploring the alignment of PhD core competencies, assessments and training requirements with students' needs. In this essay, we focus specifically on exploring how US doctoral training programs consider and use QEs during the course of PhD training. Although the terminology used to refer to QEs varies widely by discipline, department and university (e.g. candidacy, comprehensive, cumulative, general, preliminary or QE), we will refer to these as QE throughout. There is notable variability in QE formats, but consistently, QEs are executed as high stakes assessments typically administered midway through a PhD program after coursework is complete to assess readiness for dissertation research. As these are typically high-stakes exams, it seems particularly critical that QEs align with core competencies defined by PhD

programs. Here, we explore the current state of QEs and address specific questions around the purpose, perceptions and format of QEs from a survey of the peer-reviewed publications and from website analysis of peer institutes. We specifically evaluated evidence for QE format, student experience and outcomes and administration and policy. Our investigation exposes a research gap concerning the purpose, format and effectiveness of STEM QEs; reveals a lack of QE consistency across PhD programs; and identifies the need for robust investigation of QEs and their role as high-stakes exam used to evaluate PhD students' advancement to candidacy.

Methods

This mixed-methods study was conducted in two phases using a combination of qualitative and quantitative data. To investigate evidence for best practices, we first conducted a review of the extant literature (Phase 1). A scoping review was selected because the QE literature exhibits a complex and heterogeneous nature that is not amenable to a more precise systematic review. In addition, scoping review aligns with the goals of the review, which were to clarify working definitions and conceptual boundaries of QE exams and identify gaps in existing research (Peters *et al.*, 2015).

The literature review was initially focused on biomedical and basic science PhD programs. However, the paucity of reports on QEs led us to expand to any publication that discussed qualifying and related exams in doctorate education with a focus on most recent publications. A search of the literature review was conducted by using scholarly databases, including ERIC and PubMed, with a combination of search terms including “qualifying exam,” “doctoral exam,” “preliminary exam,” “candidacy exam,” “comprehensive exam” and “cumulative exam.” Articles were identified through the search by multiple researchers, duplicates were removed and abstracts were reviewed. Articles were excluded if they were:

- not focused on graduate education;
- did not address assessment of doctoral students;
- did not discuss QEs; and
- were not peer reviewed.

Papers that passed the abstract review, and a few publications that were not from peer-reviewed sources, were screened by one researcher, who independently coded the studies using thematic analysis.

To examine current policies, we then reviewed the websites of several University Graduate Schools to determine university-wide QE policies (Phase 2). We chose programs that have been identified by institutional leaders as peers of the University of North Carolina (<https://oira.unc.edu/oira/unc-system-defined-peer-group/>) and peers of the UNC Eshleman School of Pharmacy (<https://pharmacy.unc.edu/about/oe/factbook-and-benchmarks/peer-institutions/>). We examined the websites of the following universities: Duke University, Emory University, Johns Hopkins University, Northwestern University, The Ohio State University, University of Arizona, University of California-Berkeley, University of California-Los Angeles, University of California-San Francisco, University of Florida, University of Illinois at Urbana-Champaign, University of Illinois-Chicago, University of Kentucky, University of Michigan-Ann Arbor, University of Minnesota, University of North Carolina at Chapel Hill, University of Pittsburgh-Pittsburgh, University of Texas-Austin, University of Virginia- Main Campus, University of Washington-Seattle, University of Wisconsin-Madison, University of Kansas, Vanderbilt University. We manually examined Graduate School websites for examination milestones required before the thesis defense and

compared the examination descriptions, characteristics and regulations. Relevant data from the websites and policy documents were extracted into an Excel file. The Excel file included all policies related to QEs that the research team could identify on the websites, such as “Graduate School Requires QE,” “Description of QE Requirements,” “Structure of QE” and “Process for QE Evaluation.” Data were aggregated using descriptive statistics (e.g. frequencies) and thematic analysis (e.g. common patterns in text-based data). Analyses were conducted by one researcher, with audits conducted by a second researcher to promote trustworthiness. Agreement between researchers exceeded 90%.

Results: literature review

PhD training often involves a high stakes assessment before “candidacy,” which we will consider as the time when the only remaining formal requirement is completing the dissertation research, preparing the dissertation and defending the thesis. Most US and Canadian PhD programs include a QE to evaluate students before advancing to candidacy, whereas most European and Australian PhD programs require a master’s degree for entry into PhD programs in lieu of QEs (Barnett *et al.*, 2017; Bernstein *et al.*, 2014). Despite their importance in US PhD training, there is little peer-reviewed literature about QEs. As a result, we included all publications addressing PhD training – including various doctoral degrees – starting from the 1960s to gain a full picture of prior work. Notably, there were very few papers that considered QEs in a biomedical or adjacent discipline PhD program (e.g. pharmaceutical sciences, biology, chemistry, biochemistry, biophysics, genetics, biomedical engineering) (Wilson *et al.*, 2018) or provided in-depth evaluation of modern QE practices in any discipline with a doctorate degree.

Qualifying exam history

To understand the origin of QEs in PhD training, we explored their history. PhD programs are rarely subject to strict regulations from accrediting bodies, so common practices appear to be a matter of historic precedence, acceptance of previous practices and periodic change after internal and or external reviews. Additionally, faculty within doctoral programs may tend to be more research-focused rather than teaching-focused, leading to less priority being placed on the pedagogy and assessments associated with the doctoral program and more emphasis on the research performance. Prior literature reviews have noted that QEs were part of the first PhDs awarded in the USA dating as far back as the late 1800s (Estrem and Lucas, 2003; Manus *et al.*, 1992). When PhD training expanded in the 1930s and again in the 1960s, universities considered ways to standardized PhD training, and QEs were incorporated as part of the standard progression in most if not all PhD programs. With the expansion of PhD training came the burdens associated with assessing PhD trainees; as a result, the 1960s reforms to streamline QEs and move them earlier in training were commonly accepted (Estrem and Lucas, 2003). These changes led to the current cadence of course work, QEs, dissertation research and final defense as nearly universal in PhD training (Bernstein *et al.*, 2014; Estrem and Lucas, 2003; Stewart-Wells and Keenan, 2020).

The purpose of qualifying exams

As QEs are a shared practice, one anticipates that programs might also have a shared purpose in using QEs to evaluate trainees. We found that many purposes for QEs have been expressed through the years, and these continue to evolve as doctoral programs adapt to training needs. An early 1980s survey of the literature (Anderson *et al.*, 1984) found that QEs were intended:

To screen out students on the basis of ability and/or knowledge; to provide a rite of passage so the student will feel the degree has been earned; to provide an opportunity for students to organize their thinking and integrate what has been learned (Anderson *et al.*, 1984).

An analysis of composition and rhetoric PhD programs from the early 2000s (Estrem and Lucas, 2003) found that programs listed over 40 stated purposes for QEs, and these could be binned into three categories: critically thinking, expert knowledge and research/teaching ability. In a survey of counseling psychology training PhD programs (Manus *et al.*, 1992), 78% of programs reported that the purpose of QEs was to assess students' synthesis and integration of knowledge as well as basic skills and abilities (Manus *et al.*, 1992). Similarly, a survey of directors of social work doctoral programs (Furstenberg and Nichols-Casebolt, 2001) found that QEs function to test "students' mastery of knowledge and skills and assisting students' learning and progress toward the dissertation" (Furstenberg and Nichols-Casebolt, 2001). In addition to evaluating students, student performance on QEs can also be used for program evaluation (Cobia *et al.*, 2005). For example, if QEs are designed around a program's stated competencies and outcomes, demonstrated mastery of QEs can assess whether students' training prior to the QE is meeting the trainees' needs.

A more recent survey of faculty in doctoral counselor education programs identified five QE purposes, ranked by participants as:

- (1) to assess lower levels of cognitive complexity;
- (2) to assess higher levels of cognitive complexity;
- (3) to prepare student for future scholarship;
- (4) to promote a beneficial learning experience; and
- (5) to maintain tradition (Kostohryz, 2016).

Other doctoral programs report similar purposes for QEs, including biomedical sciences (Wilson *et al.*, 2018), criminal justice (Pelfrey and Hague, 2000), marketing (Ponder *et al.*, 2004), political science (McMahon *et al.*, 2020), mathematics (Earl-Novell, 2006) and nursing (Mawn and Goldberg, 2012). Student perspectives include a perception that QEs are intended to "test the capability of the candidate and to weed out weak students," though this finding was based on a small sample ($n = 35$) of engineering students (Memarian *et al.*, 2019). Clearly, the stated and functional purposes of QEs are varied and multifaceted. If the intended purpose is not specifically stated and socialized with students, the vagueness creates a risk that students and faculty will have different perspectives of QEs. A recent but very small study of faculty and students in a Doctor of Education (EdD) program showed a split on the purpose of QEs as it related to students' futures, with faculty viewing the QE as important for careers in academia, whereas students saw them as important as preparation for thesis work (Guloy *et al.*, 2020).

The complex expressions of the purposes for QEs are consistent with our interpretation leading to one possible concrete, encompassing and shared purpose of QEs: to ensure students are prepared for dissertation research enroute to their PhD. The fact that a transition to "candidacy" is frequently associated with passing QEs is consistent with this interpretation. Despite a complexity of purposes, the functional necessity of QEs can be distilled to ensuring that students are ready to progress to a focus on dissertation research. The fact that QEs are considered high-stakes assessments should lead administrators to consider many questions when establishing QE criteria and expectations. While the following are not exclusive, we consider them to be some of the most pressing in determining the format and administration of QEs within a program:

- When should QEs be administered?
- How should QEs be administered?
- How should success of QEs be assessed?
- What is the student experience with QEs?
- What is the faculty experience with QEs?
- Are QE practices equitable?
- Are QEs serving student and program needs?

The last two questions are critical as there remains a shared goal of promoting diversity in STEM PhD careers; furthermore, the job market for PhDs has greatly expanded since QEs were first accepted into the training canon, meaning that training needs have evolved as well.

Qualifying exam timing

Literature reports that QEs are offered sometime between the end of the first year or when courses are completed and the start of the fourth year, depending on several variables (Barnett *et al.*, 2017; Bernstein *et al.*, 2014). One variable that impacts timing is the institutional, program or departmental policies and requirements. For example, Berkely's PhD program in Mathematics QEs include a written preliminary exam before the end of the first three semesters and an oral QE that must be attempted by the end of Year 2 (Earl-Novell, 2006), one EdD program gives the QE in the spring semester of their second year, which is the last year of course work (McBrayer *et al.*, 2020), marketing programs often have a first year QE and a second written QE due in Year 2 or 3 (Ponder *et al.*, 2004) and QEs are typically given during the second year of most nursing PhD programs (Allard *et al.*, 2021). Other possible variables that may impact timing of a particular student's completion of QEs include individual student preparedness, bureaucratic/administrative challenges like getting the assessment committee together and lack of incentives to or penalties not to adhere to established timelines.

An important historical context of note is that the timing of QEs was designed by some programs to produce an early "off ramp" for doctoral students allowing students who are not prepared for or fully vested in completing dissertation research to exit the program be "weeded out" as it has been historically described. However, high attritions rates are not typically desired in modern doctoral programs, and hence, programs generally seek to reduce or avoid inducing the potential student failure and stress with QEs that is associated with student attrition (Estrem and Lucas, 2003; McBrayer *et al.*, 2020; Wilson *et al.*, 2018). Interestingly, QE scores only weakly correlated with degree completion in one small study of an EdD program (McBrayer *et al.*, 2020), and the timing of QEs did not negatively affect obtaining candidacy up to the 11th semester in a broad study of PhD trainees (Ampaw and Jaeger, 2012). An intriguing idea is to customize the timing of QE administration based on student needs and programmatic outcomes in a way that considers competency-based learning. Related to this, a guide for competency-based assessments in biomedical PhD training has been proposed (Verderame *et al.*, 2018).

Qualifying exam format

Generally, QEs are written, oral or a combination. Written QEs can include closed-book or open-book formats; timed or open-ended; and may include options such as on-site exams, take-home exams, written reports and research proposals. Performance assessments are

typically provided by groups of faculty, often in the form of an examining committee. Oral QEs can include presentations of a written proposal that can be based in the thesis proposed thesis research, presentation of selected topics (Mawn and Goldberg, 2012) or examination of student knowledge of critical thinking on select topics (Earl-Novell, 2006).

For programs that use both written and oral QEs, they are sometimes considered distinct steps toward candidacy in terms of timing, name, and focus. For example, the written portion might be considered a comprehensive or preliminary exam and be based on coursework, while the oral QE is based on proposed research or assigned papers. However, there is a lack of consistency in the distinctions between oral and written exams, including how they referred to by various programs. We found that some programs have explored alternative approaches. For example, some political sciences PhD programs provide students several options for QEs (Schafer and Giblin, 2008). In counselor education doctorate programs, an alternative approach using student portfolios in place of comprehensive and final oral exams has been implemented (Cobia *et al.*, 2005). A similar student-centered approach using student-developed “authentic assessments” associated with standard and optional benchmarks has been implemented in a recently established EdD program in place of traditional QEs (Stewart-Wells and Keenan, 2020).

Aligned with the idea that nontraditional QEs would provide greater value than closed-book written exams, students in engineering PhD students “showed interest in having the exam be focused on the research proposal as opposed to background theory and that they would appreciate more feedback and communication with the committee” (Memarian *et al.*, 2019). One consistent theme from our review of the literature is that knowledge-based, written assessment is considered appropriate for time/location-constrained closed-book exams, whereas open-book, take-home written assessment and other alternative assessments, such as portfolios, are more appropriate to assess higher-order thinking (Kostohryz, 2016; Pelfrey and Hague, 2000; Ponder *et al.*, 2004). One concern with alternative assessments for QEs is the challenge of standardizing process (Estrem and Lucas, 2003; Stewart-Wells and Keenan, 2020). However, one common alternative is to include a thesis prospectus in the QE, which has been reported to correlate with lower attrition and shorter time to completion (Nerad and Cerny, 1993).

Student experience with qualifying exams

As a high-stakes milestone essential to the advancement to doctoral candidacy, the QE can be stressful, and strategies to understand and enhance the student experience with QEs have been reported. This is especially important in the context of diversity and inclusivity as QEs have been identified as a potential “leak in the STEM pipeline” (Wilson *et al.*, 2018). Vagueness in expectations and lack of understanding of the exam itself increases stress and anxiety significantly, which can impact performance (DiPietro *et al.*, 2010; Harding-DeKam *et al.*, 2012; Nerad and Cerny, 1993). An older report found that students overprepared as a result of the lack of clarity, suggesting that students would benefit if programs made expectations, format and other details associated with QEs clear (Nerad and Cerny, 1993). Kostohryz and colleagues reported that 22% of reviewed doctoral programs in counselor education had no written policy for evaluating questions (Kostohryz, 2016). Similarly, 97% of reviewed comprehension and rhetoric doctoral programs gave no public statement on how exams are assessed (Estrem and Lucas, 2003). Nursing doctoral students also voiced a need to better understand evaluation criteria and exam outcomes, as variable QE experiences created a feeling of injustice and during preparation for the QE students exhibit the poorest well-being (i.e. highest stress levels and lowest satisfaction with program) (Allard *et al.*, 2021). The interviewed nursing students suggested that more support would

de-stress the QE experience, such as senior student advice to help prepare for QEs as well as peer-study groups.

One approach to overcome the stress associated with QE “unknowns” is to use an external mentoring model (Williams *et al.*, 2017). This model provided evidence of the success of mentored-group participation via a national network to navigate the hidden curriculum of academic research. Similarly, mentored guidance and peer interactions around QEs could prove crucial to effectively preparing for and dealing with the stress of navigating the hidden curriculum behind QEs. The work of Williams and colleagues (2017) suggests that implementing a mentoring program can enhance the success of underrepresented groups on QEs and beyond, which is notable given the need to increase diversity of PhD researchers. Another support mechanism is to provide mock QEs for students; this approach was successfully implemented at the MD Anderson Graduate School of Biomedical Sciences to increase retention, specifically of historically underrepresented minorities (Wilson *et al.*, 2018). In addition to providing support and clarity, the format and flexibility of QEs also impact student perceptions. For example, Stewart Wells and colleagues found that students assessed using “authentic assessments,” including writing research conference proposals or grant proposals and publishing an article in a peer-reviewed journal reported positive and motivating experiences (Stewart-Wells and Keenan, 2020). Though limited evidence-based research has been completed in the biomedical sciences, broadly across disciplines, the evidence is clear: QEs are stressful, and mechanisms to decrease stress and decrease the potential for QE associated-attrition are all reported to benefit student experiences with QEs, including additional support, clarifying the QE purpose, format and assessment mechanisms and finding ways to increase flexibility of QE formats.

Qualifying exam value and equity

Few studies have tackled complex questions regarding the equity of QEs in PhD training, and we could not identify any studies that examine evidence-based associations between QEs and postgraduate outcomes. While program structure and support have been associated with attrition, the role of QE in attrition has been only briefly addressed. The most definitive studies on overall attrition in STEM fields are from the council of graduate schools, which found that students starting STEM PhD programs between 1992–1993 and 2003–2004 had 46% completion rates at seven years and 57% at ten years for all students. A separate analysis showed that ten-year completion rates for students from URM background were slightly lower, 55% for white students, 51% for Hispanic students and 47% for black students; however, the most recent CGS report on a larger set of URM students showed a 54% ten-year completion rate (Sowell *et al.*, 2015), which suggests that there might not be a substantial demographic difference in overall completion rate.

However, attrition rates remain high, differ from program to program and can in part be attributed to program environment and support. Likewise, an earlier study found similar attrition rates and determined that program culture and student fit were the most likely causes of attrition (Lovitts and Nelson, 2000). To lend further support to the thought that programs contribute to attrition, Earl-Novell found higher rates of attrition for women and URM groups despite possessing similar GPAs (Earl-Novell, 2006). It was hypothesized the higher attrition stemmed from a lack of mentorship and role models for these groups in PhD programs. Not surprisingly, one aspect of PhD programs that may have a potentially negative and highly influential impact on equity may be QEs. The high-stakes and stress-inducing QE process has been identified as a contributor to URM student attrition while noting that program support can mitigate this problem (Wilson *et al.*, 2018). It seems natural that QEs as traditionally implemented could lead to attrition, and the evidence suggests that providing support and

potentially changing the format of QEs can lower attrition rates (Allard *et al.*, 2021; Nerad and Cerny, 1993).

A separate question is whether QEs contribute to any bias in completion rates. While potential bias has been studied in programs as a whole, there is little literature exploring bias in the QEs. By contrast, a great deal of evidence has established bias in standardized exams such as the graduate record examination (GRE), which correlate with applicant demographics but not PhD program outcomes, such as first-author publication or time-to-degree completion, suggesting that these exams might inappropriately limit admittance of minority groups that historically perform less well on the exam (Hall *et al.*, 2017; Moneta-Koehler *et al.*, 2017). As a result, many programs have explored dropping GRE requirements. Hence, following a similar argument, if it were found that QEs in fact introduced bias in program advancement and attrition, they could be inadvertently harming efforts to increase diversity of the STEM PhD workforce. It is important for programs to consider any potential biases introduced by the QE process, and proactively establish mechanisms to mitigate bias and ensure equity in the QE process.

Our review of the literature shows that there is little consistency across PhD programs on QE format, assessment, timing, student support or even nomenclature. QEs appear to contribute to attrition rates and increase student stress. For STEM PhD programs specifically, we did not find substantial evidence supporting any specific QE practices, except for providing clarity to students of QE purpose and format and providing student support during QEs.

Results: analysis of peer-university websites

The literature survey revealed mostly descriptive research on QEs, with a notable absence of QE roles in biomedical, pharmaceutical sciences and adjacent PhD programs, revealing a gap in understanding evidence-based practices for QEs. To begin the process of filling in this critical gap in our understanding of QE practices in PhD training, we examined the websites of 23 universities: 22 peer universities (i.e. university- and school-defined peers) and our home university, for institutional-wide policies on QEs. We specifically evaluated policies on the prerequisites, timing, nature, established term for QEs, as well as outcomes and objectives stated for QEs. To our knowledge, this represents the first analysis of central regulations regarding QEs representative of major US research-intensive universities.

General characteristics of university-wide qualifying exam policies

Of the 23 reviewed universities (Table 1), most ($n = 20$) have a specific graduate school or college that sets policies for graduate education, while three have graduate administration delegated to discipline-specific schools/colleges. Three of the universities without a graduate school/college, Johns Hopkins, University of Pittsburgh and the University of Virginia, maintained central oversight of graduate education through the provost office that is guided by a faculty committee. Of the 20 universities with a central graduate school, nearly all ($n = 19$) explicitly state a QE requirement for the PhD, and the University of Pittsburgh lists a university-wide QE requirement despite the absence of a graduate school. This observation is consistent with reports that QEs are regularly required in US PhD programs (Barnett *et al.*, 2017). Of the three universities without a specific central QE requirement, it is stated to be or assumed to be left to individual PhD programs, and each biomedical program examined at these, including Virginia, TX and John Hopkins, did require some form of QE.

We found that all but one university that required a QE stated that the PhD students' status upon passing changed to "Candidate." Passing QEs were further commemorated at the University of Washington-Seattle, which provides a "Candidate Certificate," and

Research question	Yes ^b <i>n</i> (%)	Summary of website data collected
Does university have a graduate school or college?	20 (87%)	Two of the universities without a graduate school oversee graduate programs through the provost's office with a faculty committee providing guidance. The other university delegates oversight to individual schools and colleges without any indication of central oversight
Does university have a QE requirement?	20 (87%)	Of the 20 "yes" universities, 19 state QE requirements of the graduate schools. Universities without a requirement leave it to individual programs, and all biomedical sciences programs checked required a QE
<i>Question asked for universities with a graduate school/college (n = 20)</i>		
Is it called a QE?	7 (35%)	Other descriptors include preliminary exam (<i>n</i> = 9), doctoral exam (<i>n</i> = 2), doctoral comprehensive exam (<i>n</i> = 1), comprehensive exam (<i>n</i> = 1), candidacy exam (<i>n</i> = 1) and general exam (<i>n</i> = 1)
Is passing the QE required to advance in the program?	20 (100%)	Policies for retaking the QE vary from school to school (in case of failure)
Does student status change after passing the QE?	19 (95%)	Eighteen call the students "candidates," one calls them All But Dissertation (ABD). One provided a certificate, and one held a ceremony
Do students need to be registered for QE?	15 (75%)	
Is a thesis proposal/prospectus required?	9 (45%)	The prospectus can be part of the QE. For each, a proposal or prospectus is commonly part of QE practices, even though not explicitly required
Is the QE format described?	8 (40%)	Two use oral format, six use oral and written format
Is a time limit for QE passing described?	7 (35%)	Graduate schools require QE completion within 3 years, 4 years or 5 years

Table 1.
Characteristics of QE policies for UNC and peer institutions (*N* = 23)^a

Notes: ^aWe explored the publicly accessible websites of the universities to evaluate campus-wide policies on QEs; ^bYes indicates that the graduate school or campus-wide resource site explicitly described or stated a given policy
Source: Authors' own creation

Duke University, which has a "A Milestone Recognition Reception" to commemorate obtaining candidate status. As QEs appeared to be a requirement for the PhD at almost all universities, we anticipated that universities would have a registration requirement and found that 15 of the 20 graduate schools/colleges did have a requirement requiring students to be actively registered in for-credit coursework when the QE was taken. We did not find any universities that offered/required a specific QE course. We also explored whether a thesis proposal/prospectus was required as part of the QE. We concluded that while it was a common option, it was not typically a requirement. We did find that several graduate schools (*n* = 9) required a thesis proposal/prospectus in some form. Each of these nine graduate schools suggested that the proposal/prospectus could be considered for the QE, but none of the evaluated graduate schools *required* that the proposal/prospectus specially be part of the QE. While we did not find substantial evidence, using the thesis prospectus as part of the QE practices was reported to decrease time to completion, and this should be explored further in future studies (Nerad and Cerny, 1993).

Graduate schools referred to QEs as preliminary exams (*n* = 9), QEs (*n* = 7), candidacy exams (*n* = 1) and general exam (*n* = 1), consistent with the widespread variability in the

literature. Irrespective of nomenclature, all graduate schools clearly defined QEs as a major milestone, and 18 (90%) defined this achievement as the primary prerequisite for candidacy status. The format and timing of the QE also varied across graduate schools. Of those that described the QE format ($n = 8$, 40%), 10% of those ($n = 2$) use oral formats and 30% of those ($n = 6$) use some combination of oral and written formats. The other graduate schools ($n = 12$, 60%) either did not describe the QE format or indicated that the format was determined by each program. As mentioned above, nine graduate schools allow programs to use the prospectus as the written QE. We did not find evidence for any further regulations on the QE format, for example, in class/take home, open/close book, course review, literature review. Similarly, we did not find any evidence regulating the specific nature of the oral exam, and it was apparent that these details were left to individual programs to define. Timing of QEs was explicitly regulated by some graduate schools ($n = 7$), with these schools requiring the exam within three, four or five years of enrolling in the program. Other timing regulations we observed included graduate schools that set the timing of the QE as the quarter or three months minimum before the final defense ($n = 4$), and some schools stated that the defense had to be within five years of the QE, or it must be retaken ($n = 3$). This variability in the mandated timing of the QE was also observed in the literature (Allard *et al.*, 2021; Barnett *et al.*, 2017; Bernstein *et al.*, 2014; Earl-Novell, 2006; McBrayer *et al.*, 2020; Ponder *et al.*, 2004). Several graduate schools ($n = 12$) set requirements prior to allowing students to sit for the QEs such as certain number of credit hours, completing required courses and being in good standing as defined by no incomplete or failing grades.

Because QEs are a major gatekeeper of PhD student progression, it is critical to consider policies QEs assessment, opportunities for remediation and program alignment. Each graduate school stated that QEs were to be assessed by a faculty committee. Graduate schools varied on whether the QE committee and dissertation committee were the same, with five schools stating the committees are the same, and two stating they are different but can overlap. Other graduate schools appear to leave this to the individual PhD programs. QE outcomes were established via the following common committee voting norms: as unanimous decision of committee ($n = 3$), only up to one “failure” vote ($n = 5$) and majority vote to pass ($n = 4$). For the other schools ($n = 4$), it was left to the individual programs to set the policy to determine passing. Generally, the QE committee also determines if a retake will be allowed, and all schools stated that an opportunity for remediation should be considered with a variety of school-specific policies regulating a possible retake.

Regarding program alignment, several graduate schools ($n = 11$) indicated a relationship of the QE to program outcomes. There was a broad range of details in these sections, with some descriptions being quite general, for example, indicating the QE should “provide written evidence of competency in the field,” and others providing a more comprehensive description of the intent of the QE by expanding on specific competencies, for example, mastery of the literature and critical thinking, expected of PhD trainees. Only the University of Minnesota indicated that programs must “Establish and communicate the program-specific purpose, format, and evaluation of the doctoral preliminary written examination” and provides a guide to help programs evaluate the QE for program alignment.

Limitations

This study has several noteworthy limitations. First, the focus of the study is on QEs, which are only one critical aspect of graduate training. As this body of research is developed, efforts should be made to understand the impact of QEs in relation to other graduate student

assessments and training experiences. Second, the data for this study was sourced from publicly available information (i.e. research, websites); while these types of data are generally intended to be accurate, they may be outdated or incomplete. Subsequent research should collect data from additional sources that can shed light on QEs. Third, the literature reviews in this study was primarily focused on and situated within specific disciplines, and the graduate schools examined were from specific types of institutions, leaving QE practices within other disciplines and contexts unexplored. While we expect that there are shared best practices for QEs, it is likely that some aspects will be discipline-specific. Finally, the website evaluation focused on university-wide policies preventing more detailed analysis of QE formats, outcomes and other critical details.

Conclusions and future directions

It is clear from our review of the literature and peer websites that QE exams have a wide variety of names, formats and regulations. Despite these variances, there appears a shared overarching purpose: a high-stakes assessment that doctoral students complete as a prerequisite to fully engaging independent dissertation research. In our evaluation, we identified some QE processes that evidence suggests will enhance student outcomes. Specifically, creating transparency in QE practices with explicit expectations and building structured student support are steps programs can take to improve QE outcomes, and we call for more research to help establish evidence-based best practices in QEs for doctoral training.

Future work should develop best practices for QE format and assessment, evaluate the relationship between QEs and student outcomes and evaluate the basic need for QEs in PhD training. In addition, the functional purposes that QEs serve should be better explored to better understand whether the QE evaluation methods are in fact predictive of any outcomes and future professional success. We propose that addressing the following topics as the basis of future studies and for programs to consider in evaluating their approach to QEs:

- identify faculty, students and administrator perceptions regarding current and best QE practices;
- assess QEs' alignment with program competencies;
- assess QEs' alignment with program needs;
- assess alignment with QEs' student needs;
- evaluate QE practices for potential bias and identify strategies for mitigation;
- evaluate effectiveness of specific QE formats;
- evaluate the impact of QE format and practices on student outcomes; and
- evaluate the need for QEs.

QEs are meant to serve as an important milestone, programmatic element and a gatekeeper to progress to advanced stages of training. By addressing these important issues with evidence-based policies and practices, there is the potential to create more effective evaluation tools, align associated outcomes with training goals, improve the student experience and create a training culture and research enterprise with greater equity. While these may seem like lofty goals, effectively preparing the future biomedical workforce is worth the effort, and an empirical approach to QEs will go a long way toward achieving progress across these priorities.

References

- Allard, E., Leclerc-Loiselle, J., Lavallée, A., Vinette, B. and Larue, C. (2021), "PhD nursing students' changing needs relative to the comprehensive doctoral examination", *Nurse Education in Practice*, Vol. 50, p. 102917, doi: [10.1016/j.nepr.2020.102917](https://doi.org/10.1016/j.nepr.2020.102917).
- Ampaw, F.D. and Jaeger, A.J. (2012), "Completing the three stages of doctoral education: an event history analysis", *Research in Higher Education*, Vol. 53 No. 6, pp. 640-660, doi: [10.1007/s11162-011-9250-3](https://doi.org/10.1007/s11162-011-9250-3).
- Anderson, W.P., Krauskopf, C.J., Rogers, M.E., Neal, G.W., Rogers, M.E. and Neal, G.W. (1984), "Reasons for comprehensive examinations: a re-evaluation", *Teaching of Psychology*, Vol. 11 No. 2, pp. 78-82, doi: [10.1207/s15328023top1102_4](https://doi.org/10.1207/s15328023top1102_4).
- Austin, A.E. (2002), "Preparing the next generation of faculty", *The Journal of Higher Education*, Vol. 73 No. 1, pp. 94-122, doi: [10.1080/00221546.2002.11777132](https://doi.org/10.1080/00221546.2002.11777132).
- Barnett, J.V., Harris, R.A. and Mulvany, M.J. (2017), "A comparison of best practices for doctoral training in Europe and North America", *FEBS Open Bio*, Vol. 7 No. 10, pp. 1444-1452, doi: [10.1002/2211-5463.12305](https://doi.org/10.1002/2211-5463.12305).
- Bernstein, B.L., Evans, B., Fyffe, J., Halai, N., Hall, F.L., Jensen, H.S., Marsh, H. and Ortega, S. (2014), "The continuing evolution of the research doctorate", *Globalization and Its Impacts on the Quality of PhD Education*, Sense Publishers, Rotterdam, pp. 5-30, doi: [10.1007/978-94-6209-569-4_2](https://doi.org/10.1007/978-94-6209-569-4_2).
- Cobia, D.C., Carney, J.S., Buckhalt, J.A., Middleton, R.A., Shannon, D.M., Trippany, R. and Kunkel, E. (2005), "The doctoral portfolio: centerpiece of a comprehensive system of evaluation", *Counselor Education and Supervision*, Vol. 44 No. 4, pp. 242-254, doi: [10.1002/j.1556-6978.2005.tb01753.x](https://doi.org/10.1002/j.1556-6978.2005.tb01753.x).
- DiPietro, J.C., Drexler, W., Kennedy, K., Buraphadeja, V., Liu, F. and Dawson, K. (2010), "Using wikis to collaboratively prepare for qualifying examinations", *TechTrends*, Vol. 54 No. 1, pp. 25-32, doi: [10.1007/s11528-009-0360-0](https://doi.org/10.1007/s11528-009-0360-0).
- Earl-Novell, S. (2006), "Determining the extent to which program structure features and integration mechanisms facilitate or impede doctoral student persistence in mathematics", *International Journal of Doctoral Studies*, Vol. 1, pp. 45-57, doi: [10.28945/60](https://doi.org/10.28945/60).
- Estrem, H. and Lucas, B.E. (2003), "Embedded traditions, uneven reform: the place of the comprehensive exam in composition and rhetoric PhD programs", *Rhetoric Review*, Vol. 22 No. 4, pp. 396-416, doi: [10.1207/S15327981RR2204_4](https://doi.org/10.1207/S15327981RR2204_4).
- Furstenberg, A.-L. and Nichols-Casebolt, A. (2001), "Hurdle or building block", *Journal of Teaching in Social Work*, Vol. 21 Nos 1/2, pp. 19-37, doi: [10.1300/J067v21n01_03](https://doi.org/10.1300/J067v21n01_03).
- Goldman, C.A. and Massy, W.F. (2000), *The PhD Factory: Training and Employment of Science and Engineering Doctorates in the United States*, Anker Pub Co, Bolton, MA.
- Guloy, S., Hum, G. and O'Neill, D.K. (2020), "Voices at the gate: faculty members' and students' differing perspectives on the purposes of the PhD comprehensive examination", *Assessment and Evaluation in Higher Education*, Vol. 45 No. 2, pp. 279-291, doi: [10.1080/02602938.2019.1637819](https://doi.org/10.1080/02602938.2019.1637819).
- Hall, J.D., O'Connell, A.B. and Cook, J.G. (2017), "Predictors of student productivity in biomedical graduate school applications", *Plos One*, Vol. 12 No. 1, p. e0169121, doi: [10.1371/journal.pone.0169121](https://doi.org/10.1371/journal.pone.0169121).
- Harding-DeKam, J.L., Hamilton, B. and Loyd, S. (2012), "The hidden curriculum of doctoral advising", *NACADA Journal*, Vol. 32 No. 2, pp. 5-16, doi: [10.12930/0271-9517-32.2.5](https://doi.org/10.12930/0271-9517-32.2.5).
- Hartnett, R.T. and Katz, J. (1977), "The education of graduate students", *The Journal of Higher Education*, Vol. 48 No. 6, p. 646, doi: [10.2307/1979010](https://doi.org/10.2307/1979010).
- Kostohryz, K. (2016), "The doctoral comprehensive examination in counselor education: faculty members' perception of its purposes", *The Journal of Counselor Preparation and Supervision*, Vol. 8, doi: [10.7729/83.1068](https://doi.org/10.7729/83.1068).
- Lovitts, B.E. and Nelson, C. (2000), "The hidden crisis in graduate education: attrition from Ph.D. Programs", *Academe*, Vol. 86 No. 6, p. 44, doi: [10.2307/40251951](https://doi.org/10.2307/40251951).

- McBrayer, J.S., Tolman, S. and Fallon, K. (2020), "Doctoral candidacy examination scores and time to degree completion", *International Journal of Doctoral Studies*, Vol. 15, pp. 181-198, doi: [10.28945/4529](https://doi.org/10.28945/4529).
- McLaughlin, J.E., Minshew, L.M., Gonzalez, D., Lamb, K., Klus, N.J., Aubé, J., Cox, W. and Brouwer, K.L.R. (2019), "Can they imagine the future? A qualitative study exploring the skills employers seek in pharmaceutical sciences doctoral graduates", *Plos One*, Vol. 14 No. 9, p. e0222422, doi: [10.1371/journal.pone.0222422](https://doi.org/10.1371/journal.pone.0222422).
- McMahon, N., Alcantara, C. and Stephenson, L.B. (2020), "The qualifying field exam: what is it good for?", *PS: Political Science and Politics*, Vol. 53 No. 1, pp. 94-99, doi: [10.1017/S1049096519001100](https://doi.org/10.1017/S1049096519001100).
- Manus, M.B., Bowden, M.G. and Dowd, E.T. (1992), "The purpose, philosophy, content, and structure of doctoral comprehensive/qualifying exams", *The Counseling Psychologist*, Vol. 20 No. 4, pp. 677-688, doi: [10.1177/0011000092204011](https://doi.org/10.1177/0011000092204011).
- Mawn, B.E. and Goldberg, S. (2012), "Trends in the nursing doctoral comprehensive examination process: a national survey", *Journal of Professional Nursing*, Vol. 28 No. 3, pp. 156-162, doi: [10.1016/j.profnurs.2011.11.013](https://doi.org/10.1016/j.profnurs.2011.11.013).
- Memarian, B., Gollish, S., Moozeh, K., Gerrard, D., Tihanyi, D. and Costa, S. (2019), "Comprehensive exam variations and experiences of doctoral candidates in engineering", in *2019 ASEE Annual Conference and Exposition Proceedings, Presented at the 2019 ASEE Annual Conference and Exposition, ASEE Conferences*, doi: [10.18260/1-2-32530](https://doi.org/10.18260/1-2-32530).
- Moneta-Koehler, L., Brown, A.M., Petrie, K.A., Evans, B.J. and Chalkley, R. (2017), "The limitations of the GRE in predicting success in biomedical graduate school", *Plos One*, Vol. 12 No. 1, p. e0166742, doi: [10.1371/journal.pone.0166742](https://doi.org/10.1371/journal.pone.0166742).
- National Academies of Sciences, Engineering, and Medicine (2018), *Graduate STEM Education for the 21st Century*, National Academies Press, Washington, DC, doi: [10.17226/25038](https://doi.org/10.17226/25038).
- Nerad, M. and Cerny, J. (1993), "From facts to action: expanding the graduate division's educational role", *New Directions for Institutional Research*, Vol. 1993 No. 80, pp. 27-39, doi: [10.1002/ir.37019938005](https://doi.org/10.1002/ir.37019938005).
- Pelfrey, W.V. and Hague, J.L. (2000), "Examining the comprehensive examination: meeting educational program objectives", *Journal of Criminal Justice Education*, Vol. 11 No. 1, pp. 167-177, doi: [10.1080/10511250000084831](https://doi.org/10.1080/10511250000084831).
- Peters, M.D.J., Godfrey, C.M., Khalil, H., McInerney, P., Parker, D. and Soares, C.B. (2015), "Guidance for conducting systematic scoping reviews", *International Journal of Evidence-Based Healthcare*, Vol. 13 No. 3, pp. 141-146, doi: [10.1097/XEB.0000000000000050](https://doi.org/10.1097/XEB.0000000000000050).
- Ponder, N., Beatty, S.E. and Foxx, W. (2004), "Doctoral comprehensive exams in marketing: current practices and emerging perspectives", *Journal of Marketing Education*, Vol. 26 No. 3, pp. 226-235, doi: [10.1177/0273475304268778](https://doi.org/10.1177/0273475304268778).
- Schafer, J.A. and Giblin, M.J. (2008), "Doctoral comprehensive exams: standardization, customization, and everywhere in between", *Journal of Criminal Justice Education*, Vol. 19 No. 2, pp. 275-289, doi: [10.1080/10511250802137648](https://doi.org/10.1080/10511250802137648).
- Sinche, M., Layton, R.L., Brandt, P.D., O'Connell, A.B., Hall, J.D., Freeman, A.M., Harrell, J.R., Cook, J.G. and Brennwald, P.J. (2017), "An evidence-based evaluation of transferrable skills and job satisfaction for science PhDs", *Plos One*, Vol. 12 No. 9, p. e0185023, doi: [10.1371/journal.pone.0185023](https://doi.org/10.1371/journal.pone.0185023).
- Sowell, R., Allum, J. and Okahana, H. (2015), *Doctoral Initiative on Minority Attrition and Completion*, Council of Graduate Schools, Washington, DC.
- Stewart-Wells, A.G. and Keenan, K.M. (2020), "Assessing doctoral students: a background on comprehensive and authentic assessments", *The Journal of Continuing Higher Education*, Vol. 68 No. 2, pp. 84-100, doi: [10.1080/07377363.2020.1743950](https://doi.org/10.1080/07377363.2020.1743950).
- Verderame, M.F., Freedman, V.H., Kozlowski, L.M. and McCormack, W.T. (2018), "Competency-based assessment for the training of PhD students and early-career scientists", *eLife*, Vol. 7, doi: [10.7554/eLife.34801](https://doi.org/10.7554/eLife.34801).

- Walker, G.E., Golde, C.M., Jones, L., Conklin Bueschel, A. and Hutchings, P. (2008), *The Formation of Scholars: Rethinking Doctoral Education for the Twenty-First Century*, 1st ed., Jossey-Bass, San Francisco, CA.
- Williams, S.N., Thakore, B.K. and McGee, R. (2017), "Providing social support for underrepresented racial and ethnic minority PhD students in the biomedical sciences: a career coaching model", *CBE—Life Sciences Education*, Vol. 16 No. 4, doi: [10.1187/cbe.17-01-0021](https://doi.org/10.1187/cbe.17-01-0021).
- Wilson, M.A., DePass, A. and Bean, A.J. (2018), "Institutional interventions that remove barriers to recruit and retain diverse biomedical PhD students", *CBE—Life Sciences Education*, Vol. 17 No. 2, p. ar27, doi: [10.1187/cbe.17-09-0210](https://doi.org/10.1187/cbe.17-09-0210).

Corresponding author

Michael B. Jarstfer can be contacted at: jarstfer@unc.edu

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com