Investigating the assessment landscape of physics graduate programs

Bill Bridges, Caitlin Solis, Joshua Barron, and James T. Laverty
Department of Physics, Kansas State University

Jacquelyn J. Chini
Department of Physics, University of Central Florida

Rachel Henderson
Department of Physics, Michigan State University

As physics graduate programs adapt to an ever-changing world, it is important to update their practices of assessing students. We are interested in the variety of ways that physics graduate students' comprehension and progress are evaluated. We conducted a landscape study of university handbooks and websites documenting the ways in which students are assessed throughout their program. These practices were compared among departments to determine how diverse department assessments are. We also compared our findings with recommendations from "Graduate STEM Education for the 21st Century" determined by a committee of the National Academies of Science, Engineering, and Medicine. This work will help to better understand what the numerous institutions across the country consider as necessary practices and requirements for graduate students. Understanding this landscape can provide a resource for graduate programs looking to update their practices and a foundation for further investigations into graduate education within physics programs.

2023 PERC Proceedings edited by Jones, Ryan, and Pawl; Peer-reviewed, doi.org/10.1119/perc.2023.pr.Bridges
Published by the American Association of Physics Teachers under a Creative Commons Attribution 4.0 license.
Further distribution must maintain the cover page and attribution to the article's authors.

I. INTRODUCTION

Recent work on graduate physics programs has focused on admissions practices. These include the reliability of the physics GRE upon admission as a predictor of success and how rubric-based holistic review contributes to a more diverse population of graduate students. [1–5]. There has been less work investigating practices within physics graduate programs beyond admissions. We know very little about the overall landscape of assessment practices required for a graduate program in physics. Ideally, these assessment practices represent what graduate programs of a university value, such as independent scholarly activity, strong research skills, and writing ability.

In addition, as physics graduate programs face an ever-changing world, they may consider updating their practices of assessing students. This could be done in the hope of addressing concerns from broadening participation of graduate students to updating practices that have not been addressed in many years. Universities expressing interest in this will likely search for common practices in other programs to model their updates after; however, we've identified that there is a lack of common understanding of the potentially varying practices of physics graduate programs across the country.

For this study, we explore formal assessment practices through publicly available graduate handbooks and websites of departments. We define formal assessment as any way that physics graduate students' comprehension and progress are evaluated towards meeting the criteria established by the department or university. We acknowledge that analyzing publicly available data is a limitation of this study. If a practice is not specifically mentioned in the website or handbook, that does not mean the practice does not exist in any capacity. Any publicly available, written documents may not fully describe the institution's assessment practices.

With a goal of better understanding the landscape of the various formal assessment practices for physics graduate programs, we ground our work in how well these practices align with the National Academies "Graduate STEM Education in the 21st Century" report [6]. This report outlines the expectations and guidelines of STEM graduate programs. In this study, we report preliminary results that address the following research questions:

- What are common assessment practices across physics graduate programs?
- How uniform are these practices across programs?
- How well-aligned are graduate student assessment practices with the recommended expectations and guidelines from the National Academies report?

To accomplish this, in Sec. II we describe the National Academies report more concretely in order to contextual our study. In Section III, we describe the methods of developing the emergent categories of assessment practices. We then report on how uniform those categories are among physics departments in Section IV. We also report on how well these emergent categories assess the recommendations of the National Academy's report in Section V B. Finally, we connect the preliminary results from this study with the numerous avenues of future work this study has to offer in Section VI

II. GRADUATE STEM EDUCATION IN THE 21ST CENTURY

The National Academies of Science, Medicine, and Engineering ordered an analysis of graduate education in STEM done in 2015 [6]. This report was completed to determine what an ideal graduate STEM education involves for all stakeholders including graduate students, faculty, and programs. It proposes recommendations and core elements of each aspect of graduate education including the Master's and Ph.D. degrees.

For the purposes of this paper we will focus on the recommendations for the Doctoral degree which include core competencies, career explorations and preparation for Ph.D students, and structure of doctoral research activities. This list of recommendations, listed in Table II, then provides a useful metric to determine how well the assessment practices of different physics departments match with what the National Academies present as important factors necessary to graduate education. For example, the report cites a core element of a quality Ph.D. education based in scientific literacy, communication, and professional skills as the learning of ethics and norms of scientific enterprise. Then when looking through our data, we look for assessment practices aligning with the idea of ethics training and norms.

III. METHODS

A. Context

The American Institute of Physics (AIP) produced a report in 2020 stating that there are a total of 260 universities and organizations that grant graduate degrees in the field of Physics in the United States [7]. This includes Ph.D., Master's in Arts or Science, Terminal Master's, Dual Credit programs, and specialty degrees or certification programs specifically in the areas of pure physics, astrophysics or astronomy, medical physics, or biophysics. As we are concerned with the general landscape of physics graduate programs, this provides a robust list of programs that include those different degrees.

B. Data Collection

For this preliminary study, we began with 60 randomly selected universities from the list provided by AIP including the home universities of the authors. We focused on data that is publicly available. A primary source we commonly used to gather information was the department handbooks of graduate programs when available. In the absence of a handbook or

if the handbook's information did not appear satisfactory, data was collected from the department websites or handbooks of the graduate college. Based on the definition of formal assessment above, we scaled down the information present for each of the university handbook or website into the distinct categories described below. For example, we collected information on practices such as the dissertation and courses required but not information covering the tuition or insurance of graduate students.

C. Data Analysis

Two researchers extracted data for each university to establish confidence. From the information gathered we determined categories of data including Courses (core and elective), Candidacy Exams (subject exams and research proposals), Dissertation, Training, and Miscellaneous requirements. The authors, spanning three universities, met to discuss category development based on the prevalence of language present in the data. These categories and sub-categories are shown in Table I. The subcategories were then determined by the most common practices that fit within each larger category. For example, if a program lists a course titled Solar System Physics it was placed in the Other Courses sub-category. In the event of multiple universities placing this course as a core course then Solar System Physics itself becomes a subcategory.

TABLE I. The list of emergent subcategories in all 5 categories from the current data of 37 universities including the number of universities that include each subcategory assessment practice.

Category	Subcategory	Num.
Courses (Core)	Classical Mechanics	26
	Quantum Mechanics	37
	Electrodynamics	32
	Statistical Mechanics	34
	Mathematical Methods of Physics	16
	Other Courses	21
Courses (Electives)	Number Required (Courses/Credits)	16
Candidacy (Subject Exams)	Written Portion	16
	Oral Portion	12
Candidacy (Research Proposal)	Written Portion	29
	Timeline	18
Dissertation	Written Portion	24
	Oral Defense	21
	Research Credits Required	9
Training	Human Subjects	2
	Learning/Teaching Assistants	8
	Responsible Conduct of Research/Ethics	6
Miscellaneous	Committees	20
	Annual Reports	4
	Journal Articles	3
	TA Requirements	8

IV. RESULTS

Due to implications of different degree programs and semester timelines, we further limited the 60 universities investigated so far to include only physics Ph.D. degree programs in universities operating on semester timelines. Doing so resulted in retaining 37 out of 60 universities. The data for these 37 universities included 19 physics graduate handbooks and 18 university websites. The list of emergent categories are listed in Table I

The following sections describe the categories, including coding decisions and subcategories when appropriate. In all cases, the number of departments that have explicit language of each subcategory is listed in Table I. The categories include: courses, candidacy exams, dissertation, training modules, and miscellaneous. Courses include the core course requirements and electives when appropriate. Candidacy Exams are split between subject exams and research proposals. The dissertation category covers the writing, presenting, and expectations of the document itself. Training covers any training the department requires students to complete, and the miscellaneous category covers any assessment that does not clearly fit within the other categories.

1. Courses

We divided the course requirements into 2 main subcategories, core courses and elective courses. Each core course was counted once if the program required students to take at least one subject. For example, if a program required a student to take two semesters of quantum mechanics we record it as one count of the department requiring quantum mechanics rather than two counts. The courses that appeared most often in the data formed the subcategories. Single instances of required courses that were common to one university were placed in the Other Courses sub-category.

For elective courses, there were two trends common in the data with programs requiring either a certain amount of courses or credits. For the sake of this study we count each university only once if they include a requirement for students to enroll in any elective course.

2. Candidacy Exams

Subject exams were considered as comprehensive examinations of core subjects offered by the department. These were written and oral tests examining the student's knowledge on specific or cumulative subjects, such as Electrodynamics and Classical Mechanics. Universities that required any form of written or oral subject exam were counted, including the subjects covered in those exams when applicable.

The more common process of achieving Ph.D. candidacy is through a student's research proposal. We considered any requirement of a student to present current research and propose a research plan. We counted universities that made any explicit mention of a student presenting research to achieve Ph.D. candidacy, regardless of whether those students also needed to pass a specified number of subject exams.

From the data we see a strong emphasis towards the research proposal as expected. While there are a significant number of departments that require subject exams as part of the candidacy process, the number was not as high as anticipated with many departments offering ways to either bypass or fulfill that requirement in different ways.

3. Dissertation

In this category we counted departments that made any explicit mention of the writing of a dissertation and defending it through a formal presentation. We considered also any mention of students being required to enroll in research credits to finish their dissertation.

We see that many departments mention the need for students to complete a dissertation of original research, yet there is a lack of more information given on what requirements are fulfilled. Something that was not as widespread was for some departments to consider the credit hours devoted to work on the dissertation.

4. Training Modules and Courses

From the data we determined that students may have to complete diverse training opportunities across institutions. We counted each mention of training requirements, including whether the training sessions were single hour modules or full semester long courses.

For example, a department requiring a student to pass a Responsible Conduct of Research (RCR) training module offered by the CITI Program would be coded as a short training module offered outside of the department.

We identified relatively few opportunities for students to be trained in different ways including human subjects, TA, or RCR modules.

5. Miscellaneous

The miscellaneous sub-categories identified so far are: committees, annual reports, journal articles, and TA (teaching assistant) requirements.

We place the faculty committees within this category as opposed to Research Proposal or Dissertation as the committees are not unique to either subcategory and likely to maintain its structure between these two milestones.

V. DISCUSSION

A. Preliminary Highlights of Analysis

From the sub-categories of core courses, we see a strong emphasis on quantum mechanics and electrodynamics. We had assumed there would be consistent requirements of the core courses and for many there are. Something notable about this data then is the lack of requirements of a classical mechanics or advanced dynamics course. Many universities in the data had this course as an option, able to be substituted by another such as Statistical Mechanics. A number of universities also required students to take a number of courses not expressed in our categories in Table I. These courses include specialty courses like optics or astrophysics, and could potentially reflect the specialities in research of that department.

We found that the most consistent language was for the writing and presenting of the dissertation. While it was the most consistent, it was also the least defined and structured of the formal practices. Many universities opted to refer to the dissertation as a piece of original research completed by the student, but not provide structure or expectations beyond this.

B. Assessing the Recommendations of the National Academies Report

There are a total of 23 recommendations from the National Academies survey in the section only pertaining to the Doctoral degree. In aligning these recommendations with the practices described in Table I, we identify 3 distinct levels: recommendations that are clearly assessed, those that may be assessed, and those that are not being assessed. Focusing then just on the group of recommendations that can clearly be assessed leaves 9 recommendations from the report. We identify the recommendations that best align with the goals of a practice listed in Table II

From this comparison, we see a strong emphasis towards the research proposal and dissertation covering many of the elements from the National Academies report. From looking at the university handbooks, there are numerous mentions of a student's dissertation but not about the type of content it has beyond being a piece of original research. The requirements for the dissertation then are left to be determined by the committee or the norms of that student's field of study. This

TABLE II. Emergent assessment categories from the data that best aligned with the recommendations of the National Academies survey.

Learning Goals	Courses	Candidacy Exams	Dissertation	Training	Miscellaneous
Students should develop deep specialized expertise	X		X		
Students should articulate an original research question		X	X		
Students should design a research strategy		X	X		X
Students should evaluate outcomes of research		X	X		
Students should adopt rigorous standards of investigation		X	X		
Students should learn and apply professional norms, ethical responsibilities of scientists, and ethical standards		X	X	X	
Students should develop the ability to work in interdisciplinary team settings		X	X		X
Students should acquire the capacity to communicate	X	X	X		X
Students should develop professional competencies	X	X	X		

ambiguity of the content and requirements of the dissertation makes it difficult to precisely align the recommendations of the National Academies survey.

Another notable result from this comparison is the mention of students learning ethical standards and norms. Of the 37 Ph.D. programs we found only six mentions of ethics training. However, since our analysis is limited to publicly available information, it is possible more universities require ethics training and communicate those requirements to students through other means.

C. Study Limitations

The results of this study are based on public data. We prioritized graduate handbooks for each university but necessarily used the websites if the handbooks were not available or proved insufficient for data collection. By using this type of data, we continually face the possibility that the data is incomplete or inconsistent with what is actually practiced at each university. As evidenced by nearly half of the Ph.D. programs not having available graduate handbooks, we do not have a consistent form of data to use in this study. In future research, we will corroborate these findings with departments.

VI. CONCLUSIONS AND FUTURE WORK

This analysis represents preliminary results from a subset of data that covers only Ph.D. degrees from universities that operate on semester timelines. We determined a list of assessment practices from this data organized into five main categories: Courses, Candidacy Exams, Dissertation, Training, and Miscellaneous. In comparing these categories among departments, we see a tendency towards many universities having explicit language about the expectations of dissertations, research proposals, and course requirements.

In investigating how well the recommendations of the National Academies report are being assessed in these departments, we found some that were clearly being well assessed and those that were not. Those being assessed are largely focused on the research proposal of the Candidacy Exams and Dissertation categories. Recommendations from the survey report that were not clearly being assessed were more focused on the Courses or Training categories.

In particular the dissertation holds a lot of ambiguity of specific recommendations and how they are directly assessing the recommendations of the National Academies survey. Investigating this ambiguity and how it connects to those recommendations lies within future work of this study.

As of now we have finished coding only 60 of the 260 total universities in our data. In finishing this coding process, we can begin the numerous comparisons between the university practices and how well all of the recommendations of the National Academies survey align with them. Future work then includes: interviewing faculty to confirm practices at varying institutions, comparing practices between all of the degree types, and the various ways that graduate programs value their students' time.

ACKNOWLEDGMENTS

We would like to thank the College of Arts and Sciences of Kansas State University for their support. Also thanks to KSUPER for their valuable insights. This material is based upon work supported by the National Science Foundation under Grant No. 2013339

[1] M. Verostek, C. W. Miller, and B. Zwickl, Analyzing admissions metrics as predictors of graduate gpa and whether graduate gpa mediates ph. d. completion, Physical Review Physics Education Research **17**, 020115 (2021).

[2] N. T. Young, N. Verboncoeur, D. C. Lam, and M. D. Caballero, Rubric-based holistic review represents a change from traditional graduate admissions approaches in physics, Physical Review Physics Education Research **19**, 010134 (2023), publisher: American Physical Society.

[3] N. J. Mikkelsen, N. T. Young, and M. D. Caballero, Investigating institutional influence on graduate program admissions by modeling physics graduate record examination cutoff scores, Physical Review Physics Education Research **17**, 010109 (2021).

[4] C. Miller and K. Stassun, A test that fails, Nature **510**, 303 (2014), number: 7504 Publisher: Nature Publishing Group.

[5] N. T. Young, K. Tollefson, R. G. Zegers, and M. D. Caballero, Rubric-based holistic review: a promising route to equitable graduate admissions in physics, Physical Review Physics Education Research **18**, 020140 (2022).

[6] A. Leshner and L. Scherer, eds., *Graduate STEM Education for the 21st Century* (National Academies Press, Washington, D.C., 2018).

[7] Roster of Physics Departments with Enrollment and Degree Data, 2020 | American Institute of Physics, https://www.aip.org/statistics/reports/roster-physics-departments-enrollment-and-degree-data-2020.