Investigating measures of graduate student retention

Bill Bridges 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

Research in graduate education has largely been focused around admission practices and retention. Within the physics community specifically, there is less information on the rate of students receiving degrees. Motivated by understanding the connection that retention has with institutional requirements, we propose three measures utilizing survey data collected for decades by the American Institute of Physics. Utilizing these measures for the 284 graduate programs, we find that the average yearly retention of graduate students is 97%, and the percentage of first year graduate students who are awarded PhDs and Master's is 57% and 27% respectively. These measures highlight a significant difficulty in establishing a yearly measure for changes in graduate student populations, but a good starting point into the rate of students earning degrees. To further this work, we must ensure that the measures can become robust enough to be applied to individual institutions.

2024 PERC Proceedings edited by Ryan, Pawl, and Zwolak; Peer-reviewed, doi.org/10.1119/perc.2024.pr.Bridges

I. INTRODUCTION

Investigating graduate education generally falls into two categories: admission practices and student retention. For the purposes of this work we will be focusing on the latter. There are numerous studies on the reasons that students leave a graduate program, motivated by results of studies such as Lovitts who concluded that the number of graduate students receiving their degrees is as low as 50% [1]. This measure has been largely done for undergraduate students, also finding that the attrition of students is nearly 50% [2, 3]. Researchers have used these measures to start looking at the experiences of students in their programs [3–8].

In shifting focus then to the retention of physics graduate students, looking at this measure is not new to education research. In 2005, a joint task force between the American Institute of Physics (AIP) and the American Association of Physics Teachers (AAPT) surveyed the state of physics graduate education. The report from this survey highlights recommendations for furthering the education of graduate students while highlighting places of concern for the community to address[9]. While reports like this one are instrumental in showing the number of graduate students enrolled as first years or how long it typically takes for students to graduate, there is still a lack of information of several aspects of graduate education.

As part of an ongoing investigation into the graduate requirements of physics students, we are interested in the potential connections between an institution's retention of graduate students and its program requirements. An exploratory step in this direction is to start investigating the changes in graduate student populations and retention rates at a national level for all physics graduate programs. In particular the research questions we aim to answer with this particular paper is:

- What is the national retention of physics graduate students?
- What is the rate of first year physics graduate students earning degrees?

To investigate this question, we analyzed over 20 years of AIP survey data (Section II) to develop several measures (Section III). We contextualize the results of these measures using the total information from all institutions (Section IV), and discuss those results (Section V) while identifying places of concern and limitations with the AIP data. Finally (Section VI) we propose avenues of further investigations into modifying these measures for individual institutional analysis.

II. AIP SURVEY DATA

The data for the entirety of this study comes from the yearly reports from the American Institute of Physics (AIP) [10]. These reports represent the enrollment and degrees granted of each institution of the United States. The information is self-reported from each institution and the report states both the total enrollment and degrees granted across the nation and from each school individually. The data collected includes institutions ranging from exclusively granting bachelor's degrees to graduate degrees and certifications in physics, applied physics, medical physics, etc.

These reports have been collected and aggregated for decades, but for the purposes of this study we focused on the reports from 2000 until 2022. The reports each average around 750 individual institutions. As we are concerned with graduate enrollment and degrees, we constrain this data set to cover institutions granting those degree types. This would then exclude all bachelor's-only granting institutions and those that offer Master's certificates or accelerated Master's degrees as those institutions may not have their own graduate programs. Doing so reduced the full data set to 284 institutions. The categories of data that apply to these institutions are described in Table I. The full data set is not complete, as 20% of the total potential data points are not being reported by institutions according to AIP.

III. DEVELOPING MEASURES

A. Yearly Graduate Student Retention

The first measure we introduce focuses on the yearly retention of graduate students, comparing the number of students entering and leaving their graduate program. The total number of students in year $n+1$ should be equal to the total number of students in the previous year (T_n), minus the students awarded PhDs (P_n) and Masters (M_n), adding in the number of First years enrolled in the given year (F_{n+1}), and minus any students who are not enrolled for other reasons (N_n).

$$T_{n+1} = T_n - P_n - M_n + F_{n+1} - N_n \tag{1}$$

As we are focused on the retention of students, with respect to the total population, we then define Retention (R_n) as:

$$R_n = 1 - \frac{N_n}{T_n} = \frac{T_{n+1} + M_n + P_n - F_{n+1}}{T_n} \tag{2}$$

B. Degrees Awarded to First Year Graduate Students

This type of measure we introduce focuses on the rates at which First Year graduate students are awarded degrees. This measure focuses on the so-called 'success' rates of students earning their degrees, and it is necessary to determine a time scale to determine this measure. A result of the AIP/AAPT task force report stated that the average time it takes for graduate students to earn a PhD in physics is 5 to 6 years, so we propose this as a suitable time frame for this scale [9]. Beginning then with a second measure that determines the rate of First Year students who are solely awarded with a PhD, we propose the following expression:

TABLE I. AIP survey data collected in Fall of each year. Each variable represents the sum of 284 institutions.

Var.	Name	Description
F_n	First Year Graduate Student Enrollment	Number of students who started Graduate School in *calendar* year n.
P_n	PhDs Awarded	Number of PhDs awarded during the *academic* year beginning in Fall of year n.
M_n	Exiting Master's Awarded	Number of terminal Master's degrees awarded during the *academic* year beginning Fall of year n.
T_n	Total Graduate Students	Number of Graduate students enrolled in Fall semester of year n, including first year and degree recipients.

$$(PA)_n^{(x)} = \frac{\sum_{i=-x}^{x} P_{n+6+i}}{\sum_{i=-x}^{x} F_{n+i}} \tag{3}$$

Here, $(PA)_n$ (PhDs Awarded) is the percentage of First Year students who are awarded a PhD six years later. While there may be an average time scale for students to be awarded a PhD, it is not inclusive to those students earning degrees outside this time frame. To account for this, we impose a scaling factor, (x), as a way to involve more data for each year's calculation. The two values of x chosen for the rest of this paper will be $x = 0$ and $x = 1$. In the case when $x = 0$, then the expression above returns a direct comparison between the First Year population at some year n and the PhDs awarded 6 years after. In the case of $x = 1$, the expression then averages each variable with the surrounding years. For example, if calculating the percentage at $n = 2010$, then the expression would average the PhDs awarded in 2015, 2016, and 2017; then compare it to the average First Year enrollments of 2009, 2010, and 2011. Here we can capture more realistic time scales for graduate students, accounting for ranges of students being awarded degrees within 4 years up to 8 years later.

Looking then at the rate of Master's degrees awarded to the First Year enrollments, this third measure is similarly constructed. As above, we need to establish a suitable time frame for this degree type. From another study from this group, many Master's students are expected to defend their theses/reports within 2 years. With similar parameters we define the Master's awarded, (MA), as:

$$(MA)_n^{(x)} = \frac{\sum_{i=-x}^{x} M_{n+2+i}}{\sum_{i=-x}^{x} F_{n+i}} \tag{4}$$

This expression operates on the same scaling structure as the PhD, where the minimum year for a degree to be awarded is under 1 year and the maximum is 4 years after the First Year enrollment.

IV. RESULTS

A. Yearly Retention of Graduate Students

The Yearly Retention measure of the total data set from all 284 institutions can be found in the second column of Table II. Using this measure, the average retention rate from 2000-2022 was 97%. The minimum retention rate is 85% and the maximum is 116%. The standard deviation for these results of this measure was 7%.

The average retention of students is below 100%, suggesting that there are a number of students not enrolled between years as expected. However there are a number of years that do not follow this trend, sometimes drastically such as an increase in population of 27 percentage points between 2008-2009 and then a drop by 30 percentage points the following year. The years that are above 100% call this measure into question. Such a value suggests that after accounting for students who graduated and those who started a program, there was an increase in the number of students enrolled. With no further information on the direct causes for the results, we find it difficult to trust these results and strongly suspect this indicates significant problems with the underlying data; specifically with the Total Graduate Students values. T_n seems to be the likely problem, as the form departments fill out is least specific about the time period for this value.

B. Rate of First Year Students Awarded a Graduate Degree

The results of applying the second measure can be found in columns three and four of Table II. In the $x = 0$ case, the average rate of first year students awarded a PhD 6 years later is 57%. The minimum rate of students awarded this degree is 46% and the maximum rate is 65%. Then in the $x = 1$ case, where the rates are averaged over 3 years, the average rate was also 57%. The minimum rate was 50% and the maximum was 61%. We find the result is comparable to the established literature of graduate students who earn PhDs being as low as 50%

The results of applying the last measure can be found in columns five and six of Table II. In the $x = 0$ case, the average rate of students awarded a Master's after 2 years is 27%. The minimum rate of students awarded a Master's is 23% and

the maximum is 31%. Then in the $x = 1$ case, where the rates are averaged over 3 years, the average rate of degrees being awarded is also 27%. The minimum rate is 24% and the maximum is 29%.

TABLE II. The results of each measure to the full data set totalling 284 institutions over 4 categories. Here n represents the yearly reports which is used as an index for each measure. R_n is the retention percentage of a year n based on the total graduate student population. $(PA)_n^{(0)}$ and $(PA)_n^{(1)}$ are the rate of First Year Graduate Students graduating with PhDs on an average of 6 years later. Similarly $(MA)_n^{(0)}$ and $(MA)_n^{(1)}$ are the rate of First Year Graduate Students graduating with Master's on an average of 2 years later. Blank spaces show when the expression would involve invalid or blank data.

n	R_n	$(PA)_n^{(0)}$	$(PA)_n^{(1)}$	$(MA)_n^{(0)}$	$(MA)_n^{(1)}$
2000	108	51		23	
2001	98	51	52	24	24
2002	95	53	50	25	25
2003	98	46	50	25	25
2004	92	51	51	25	26
2005	94	56	55	27	26
2006	94	59	58	26	27
2007	90	58	59	28	27
2008	89	59	60	27	27
2009	116	62	59	24	25
2010	86	55	60	25	25
2011	102	63	59	25	25
2012	85	58	61	26	26
2013	101	61	59	29	28
2014	105	59	58	28	28
2015	96	55	60	28	29
2016	99	65		29	28
2017	99			27	28
2018	96			29	27
2019	101			26	29
2020	97			31	
2021	98				
Average	97	57	57	27	27

V. DISCUSSIONS

A. Difficulty in a Yearly Retention Measure

In attempting to capture graduate student retention, we wanted to see how the population of graduate students would change on a yearly basis. This aspect of 'retention' would hopefully illuminate any productive avenues for further investigations at the institutional level. In applying this measure, we instead find results that are difficult to believe at this time. We designed this measure based on the varying time scale present in the AIP reports. There were a few decisions made to make sense of the different time frames, and one that is not expressly clear is in the Total Graduate Student Population. Unlike the other categories, there is not a specified time frame for this group, and in light of this ambiguity we assigned a time frame to this category that best fit the characterization of the full data set. The current categories also do not cover instances of students transferring in or out of their programs, or if those students are dropping out of their programs. An additional point of concern that could be causing the measure to fail is in the data itself. There is potentially some subtlety in the data that we are not capturing at this time that could lead to results over 100%. As such, this measure fails to capture how the global graduate student population is changing each year, and will require significant work before it can first be applied on an institutional level.

B. Degrees Awarded to First Year Graduate Students

The second half of this study is concerned with the how likely it is the First Year graduate student population will be awarded a PhD or a Master's. Placing this in the context as described above, we find this result believable as it is consistent with data from the literature [1, 8]. This then opens up a number of questions, such as: should this number be consistent with data from over 20 years ago, or is it reasonable to expect this number to be higher considering the body of work looking at improving graduate student experiences? With this set of results appearing reasonable, further concern is placed on the first measure. A noted difference in the formation of these measures is that this measure does not incorporate the Total Graduate Student T_n.

For the measure focusing on the percentage of First Year students who are awarded a Master's degree, it is difficult to clearly make sense of. As described by the AIP survey itself, the degree awarded is listed as 'Exiting Master's' and covers all students who enter their program solely for the Masters and those students who earn one in lieu of a PhD. Taking both measures together gives that the rate of a First Year graduate student to earn a graduate degree at 84% is similarly interesting.

C. Study Limitations

There are a number of concerns that come from this study. First is a lack of consistency in the full data set. We purposefully applied these measures to the total of all institutions to minimize any potential unreported data forcing the measures to be volatile. 20% of the total data points of all institutions were not reported from all 20 years of data. Another aspect of the data that is a concern is the potential ambiguity of the self-reported data from each institution. With each institution responsible for providing data for the survey, there is no guar-

antee in collecting that data consistently each year, or that the data presented is correct.

A major concern that we have that was apparent for the whole study is the way that the information for the survey was collected. The 4 categories each have an associated time frame that the data could be collected for. In developing the first measure we had many discussions of how these time frames overlap between Calendar and Academic Years, particularly with double counting students between those graduating in the same term that they are being counted within the Total category. The measure then represents our best interpretation of the varying time frames. This is particularly important for the category of Total Graduate Students, which is present in the ill-behaved first measure, and not for the more reasonable second and third measures.

In advancing this work, with only 60 institutions having complete data from 2000 to 2022, we are unsure at this time if any difficulty in applying these measures at the institutional level could be due to the unreported data or another reason. An additional concern is to ensure that the measures used are able to account for institutions of varying sizes. Within the survey data there are institutions that have a total graduate student population over 100, and others that have a population below 10. By using the total data of all 284 institutions, the large enrollment institutions may be dominating the results and any potentially interesting variances in small enrollment institutions may be lost by this scale.

VI. CONCLUSIONS AND FUTURE WORK

This work is an exploratory study into determining the feasibility of outlining measures of understanding the yearly retention of graduate student population and rates of first year physics graduate students earning a Master's or a PhD. The study incorporates data from the American Institute of Physics yearly surveys that includes the yearly enrollment and degrees earned by graduate students. Investigating data from a total of 284 institutions over 20 years, we found that by one measure an average of 97% students are staying in their graduate programs each year. By the second set of measures of the same data set we found that around 57% of first year students are awarded a PhD, and 27% for students awarded a Master's. We find the first measure to be ill-behaved and produce results that are difficult to believe. The second measure produces results that are more in line with the established literature of graduate education. The third measure additionally offers an interesting comparison to be made for future investigations. In total we find a promising start to developing a retention measure for physics graduate students, and a reasonable rate of students earning graduate degrees consistent with literature.

There are multiple opportunities for advancing this work. This group has an ongoing investigating of the general requirements of physics graduate students. In an effort to connect this paper with the larger study, the natural path forward is to adjust the measures developed here to be applied to the 284 institutions individually. The first measure requires some significant work to better capture how the retention of graduate students is changing each year before it can start to be applied on the institutional level. The last two measures are in a better position; however, care must be taken to ensure that the measure can be scale-able to the varying institution sizes. Ensuring that the data that was used for this study is better understood for the varying time scales, and that the measures can comfortably cover years of unreported data is critical for furthering this work. Connecting the changing graduate student population and retention measures with each institution's graduate requirements, we can start to compare programs. This study also provides a similar foundation to Lovitt's work, where a combination of this data set with qualitative analysis can further the understanding of current physics graduate student experiences.

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] B. E. Lovitts, *Leaving the Ivory Tower: The Causes and Consequences of Departure from Doctoral Study* (Rowman & Littlefield Publishers, Inc) ERIC Number: ED462026.

[2] J. Summerskill, Dropouts from college (1962).

[3] J. P. Bean, Dropouts and turnover: The synthesis and test of a causal model of student attrition, Research in Higher Education **12**, 155, publisher: Springer.

[4] J. P. Bean and B. S. Metzner, A conceptual model of nontraditional undergraduate student attrition, Review of Educational Research **55**, 485, publisher: [Sage Publications, Inc., American Educational Research Association].

[5] J. B. Berger and J. F. Milem, The role of student involvement and perceptions of integration in a causal model of student persistence, Research in Higher Education **40**, 641, publisher: Springer.

[6] J. M. Curtin, G. Blake, and C. Cassagnau, The climate for women graduate students in physics, Journal of Women and Minorities in Science and Engineering **3**, 10.1615/JWomenMinorScienEng.v3.i1-2.70, publisher: Begel House Inc.

[7] D. Sachmpazidi and C. Henderson, Departmental support structures for physics graduate students: Development and psychometric evaluation of a self-report instrument, Physical Review Physics Education Research **17**, 010123, publisher:

American Physical Society.
[8] N. A. Gross, R. Lopez, W. J. Hughes, and H. M. Oluseyi, Theory into practice: Applying lessons learned from retention studies to build a diverse graduate community, in *AIP Conference Proceedings*, Vol. 991 (AIP) pp. 110–118, ISSN: 0094243X.
[9] D. Campbell, Chair, T. Appelquist, R. Diehl, J. Fajans, J. D. Garcia, J. Gates, A. Goldman, P. Jung, and M. Paesler, Graduate education in physics: Report of the joint aapt-aps task force, (2006).
[10] Roster of physics departments with enrollment and degree data, 2022 - aip.org, https://ww2.aip.org/statistics/roster-of-physics-departments-with-enrollment-and-degree-data-2022.