2007 RESEARCH GRANT
FINAL PROJECT REPORT

Date: June 26, 2009

AIR Award Number: RG 07-209

Principal Investigator Name: Patricia B. Cerrito

Principal Investigator Institution: University of Louisville

Secondary Principal Investigator Name(s):

Secondary Principal Investigator Institution(s):

Award Amount:$30,000

Proposal Title: An Investigation of University Expectations of Work

1. List of current and pending publications based on the project's findings

- Book, entitled, Investigations of University Expectations Using Text Documents and the Enrollment Database. Cerrito PB. Pending with Bentham Science, Inc.
- **Cerrito PB.** A Classroom Experience of Research for Undergraduate Mathematics Majors and General Education Students. Quarterly of the Council for Undergraduate Education. January, 2008.
- **Cerrito PB.** Panel. Data Mining Tools Compared, Clementine, Enterprise Miner, YALE and Insightful Miner. Air Forum. Seattle, WA. May, 2008.
- **Cerrito PB.** Internet Investigations of Text Material to Compare Programs Across Institutions. Air Forum. Seattle, WA. May, 2008.

Presentations
Cerrito PB. Panel Presentation. Data Mining – Concepts, Myths and Case Studies. AIR07. Kansas City. June, 2007

2. Demographic information about individuals funded under the grant
   Principal Investigator Name: Patricia B. Cerrito
   Gender: F
   Race/Ethnicity: Caucasian
   Citizenship: USA
   Disability Status: None

Final Report

A Detailed Examination of Workload Assigned and Faculty Productivity: Comparison Across Departments

Abstract
The purpose of this study is to examine assigned faculty workloads in Biology and Mathematics and to compare them to faculty productivity as listed in faculty vitas. This study used publicly available records at the University of Louisville to compare and contrast the two departments. Exploratory data analysis and data mining techniques were employed to investigate the data. Results showed that Biology is more productive compared to Mathematics in terms of research and grants received. Mathematics has increased the number of administrative positions that include course release, with 6 such positions in spite of having only 25% of the students compared to Biology. Universities need to make such global comparisons between departments to optimize faculty outcomes.

Keywords: *Workload, Productivity, Quality assurance*

1. Introduction

For many years faculty workload and productivity had been studied for many reasons such as efficiency, performance or enhancing academic policies; on the other hand those studies had different approaches. The term "Faculty workload" has been defined with different prospective and viewpoints, for example it was referring for the teaching percentage, research activity or community services. In addition to administrations role to ingrate department activity and enhance overall department performances. While the term "Faculty Productivity" was used to measure what was produced on faculty time, for example number of publications, number of instructed class and external grants.

In this paper, we compare the Departments of Mathematics and Biology to see if there are any differences in the levels of workloads and productivity between them and illustrate the possible reasons for that variation. Both workload plans and curriculum vita will be investigated, In particular, external grants number, number of publication vs. allocated research time. Also analyze the impact of time allocated for administration roles on the faculty productivity. There are some troubling aspects in that faculty who have ceased to publish receive up to 25% of their time for research. In addition, many faculty moves on to administration with considerable time allocated to service. It appears that some roles have inflated percentages while others are under-valued in terms of workload.

Teaching and research represent significant portion of faculty's time which highlight the need to examine in details how faculty utilize their research time. It revealed that there is no uniform approach on how faculty utilizes his research time, for example some faculty will use it on internal or non funded research while other will focus on external grants more, Data Mining sections shows more details.

This research results based on analysis and study of public records, any personal information have been excluded to comply with privacy act.

2. Data Summary

Faculty's work plan (AWP) and curriculum vita uses two types of variable: interval and nominal, this section represents brief description the dataset. As shown on Table 1, Biology (BIO) and Mathematics (Math) departments are equally represented on the sample.

Dept				
Dept	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
Bio	52	50.00	52	50.00
Math	52	50.00	104	100.00

Table 1: Department's Frequency

For data extraction, the most recent curriculum vita has been used, as well AWPs were collect for the years (2004-2009) as shown on Table 2 the most recent years represent the largest portion of the dataset.

Year				
Year	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
2004-2005	1	0.96	1	0.96
2006-2007	19	18.27	20	19.23
2007-2008	38	36.54	58	55.77
2008-2009	46	44.23	104	100.00

Table 2: AWP Year's Frequency

Since participator's rank is important factor on faculty workload and productivity analysis, the dataset enveloped the different faculty ranks: Assistant professor, Associate professor and Professor. As shown on the following table (Table 3); the highest contribution was by Professor Rank followed by Assistant and then Associate Professor. While figure 1 shows the mass distribution of each rank over the departments.

Rank				
Rank	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
Assistant professor	38	36.54	38	36.54
Associate Professor	24	23.08	62	59.62
Professor	42	40.38	104	100.00

Table 3: Rank's Frequency

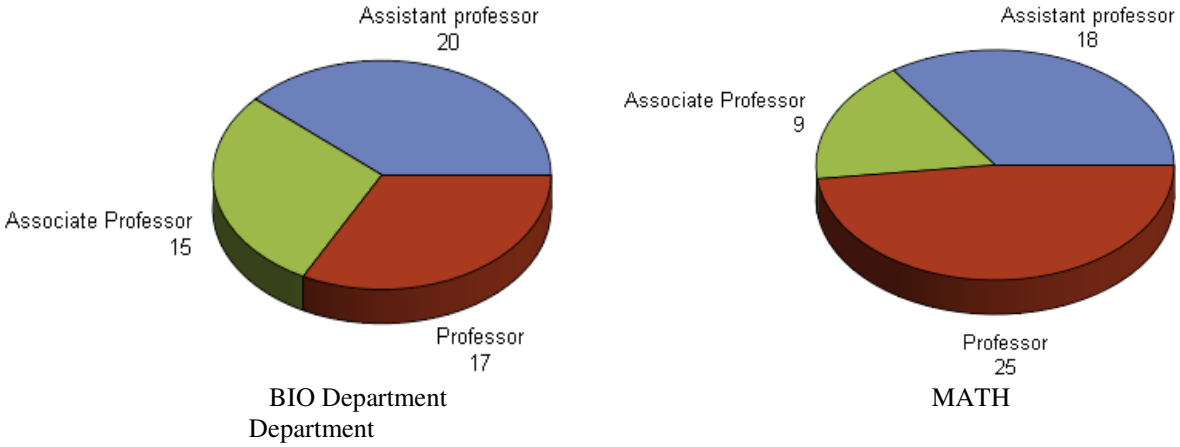


Figure 1: Faculty's Rank Frequency per Department

Another significant factor for faculty workload analysis is the number of active external grants, Table 4 shows NIH and NSF are the most popular funding organization where they represent seventy present of the current active external grants.

External Grants				
Funding Org.	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
Bayer Pharma	1	3.45	1	3.45
EPA	1	3.45	2	6.90
NIH	13	44.83	15	51.72
NSF	7	24.14	22	75.86
ORAU	1	3.45	23	79.31
Shulsky	2	6.90	25	86.21
US NPS	1	3.45	26	89.66
USDA	2	6.90	28	96.55
N/A	1	3.45	29	100.00

Table 4: Funding organization's Frequency

By comparing Mathematics and Biology departments in terms of active external grants, it was clear that Biology department has considerably more variability on funding organizations, which is not valid for mathematics department as Figure 2 and 3.

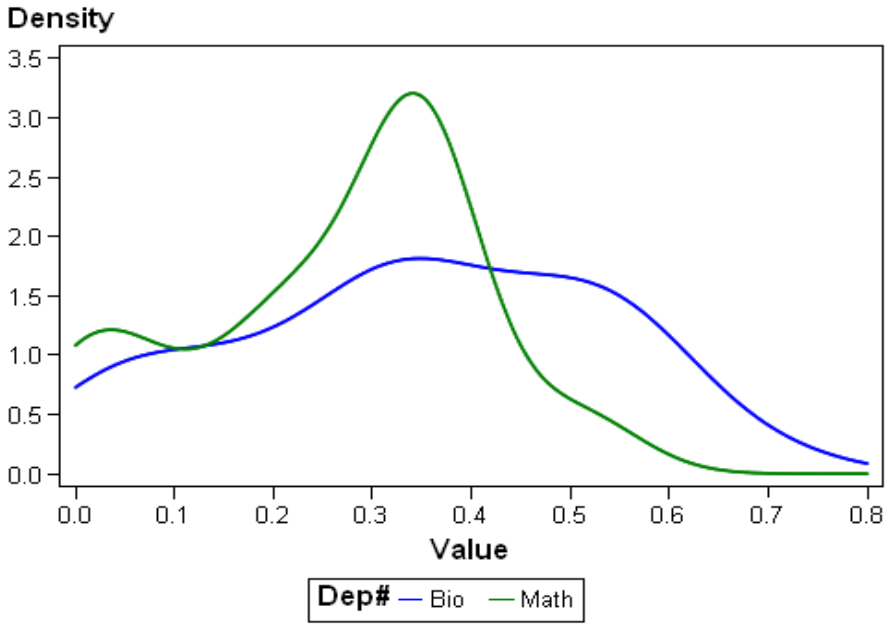


Figure 2: Assignment of Workload to Research Activities

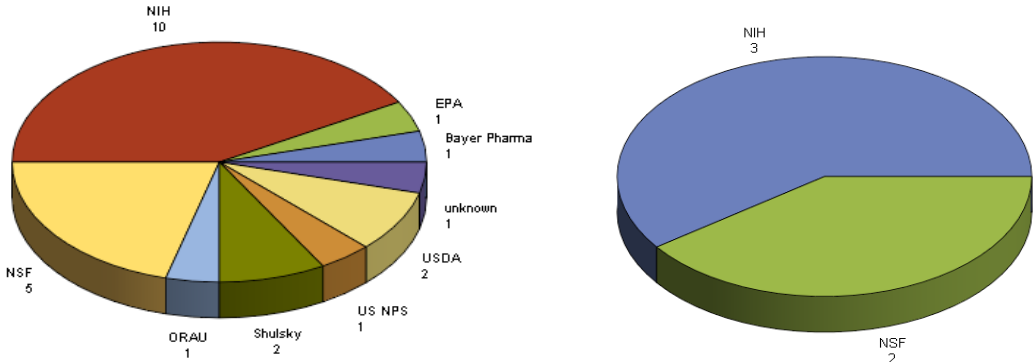


Figure 3: Active External Grants

In contrast, the research for Mathematics peaks sharply around 35%. The teaching assignment is lower for Biology compared to Mathematics (Figure 4).

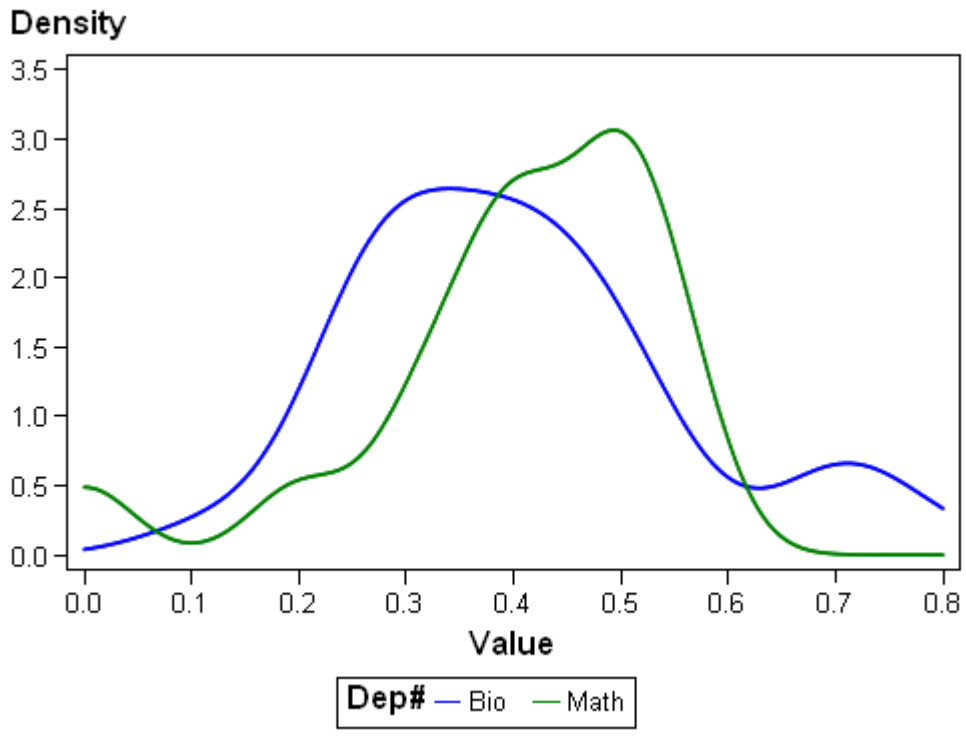


Figure 4: Assignment of Workload to Classroom Teaching Activities

Mathematics peaks at 50% while Biology peaks at 35%. The potential reason of that Biology has a lower teaching load and a more varied research load is very likely because of grants activity. Therefore, we want to examine the number of grants by department. We also want to examine teaching activities other than those in the classroom. As it turns out, the Department of Biology has a very high proportion of time for student theses and dissertations; the Department of Mathematics has a negligible allocation for this activity. Figure 5 gives the allocation for the Department of Biology. Note that the peak is 4%, with a high of 14%.

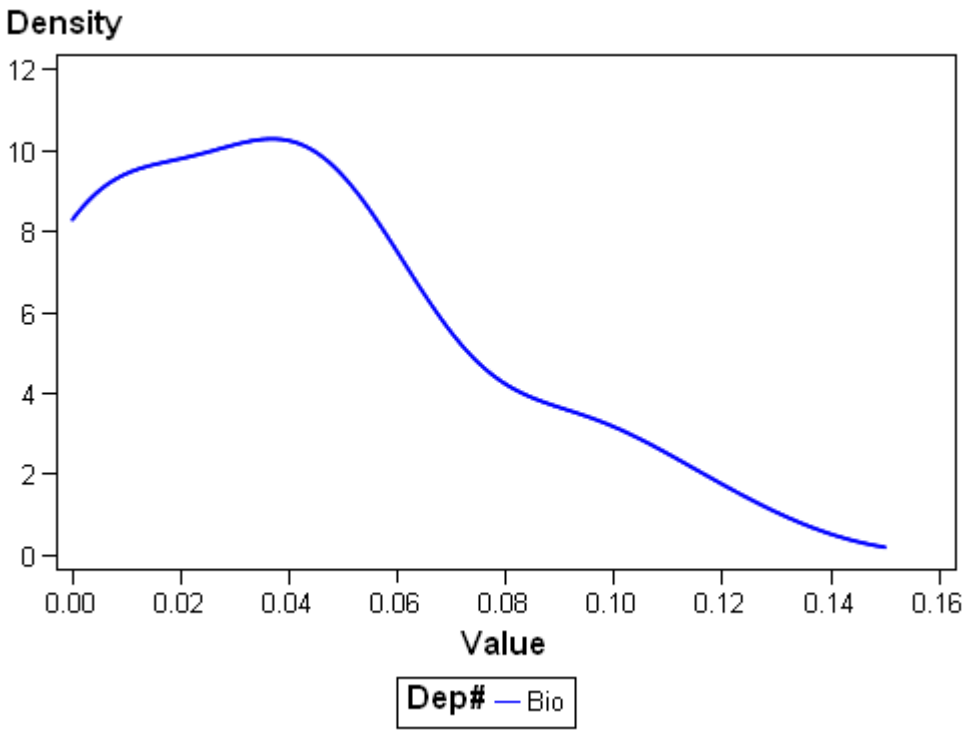


Figure 5: Time Allocated to Supervise Theses and Dissertations

For the three-year period 2006-2007, 2007-2008, 2008-2009, Biology supervised 153 thesis students allocating 82% of an FTE to do so. In the same time period, Mathematics supervised 21 students for 18% of an FTE. There is also a difference in the number of grants by department. Over the three year period, Biology has 27 grants; in the same time period, Mathematics has 7. It suggests that grants are far more highly valued in the Department of Biology compared to the Department of Mathematics. In addition the publication number for the MATH department is much less than the BIO department as show but the KDE on the following figure (figure 6).

Interestingly enough, the Department of Mathematics has 2.70 FTE's allocated for administrative activities while the Department of Biology has only 1.175. It is not clear just why the Department requires more administrative time, so we will investigate the positions in more detail.

Also, the number of publication reflect that BIO department are more productive which is typical outcome for higher number of research grants and high proportion of time for student theses and dissertations supervision.

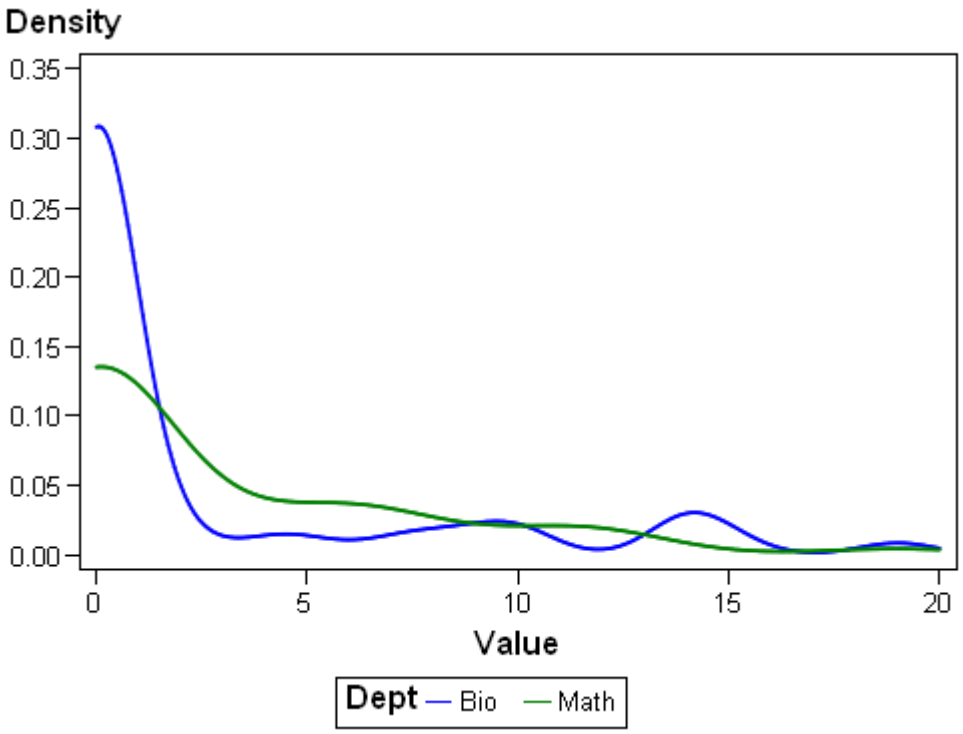


Figure 6: Publication number per department

It is clear the most of the Assistant professor population is contained in narrow area while the Professor shows wider variability as well as the Associate professor as shown on figure 7.

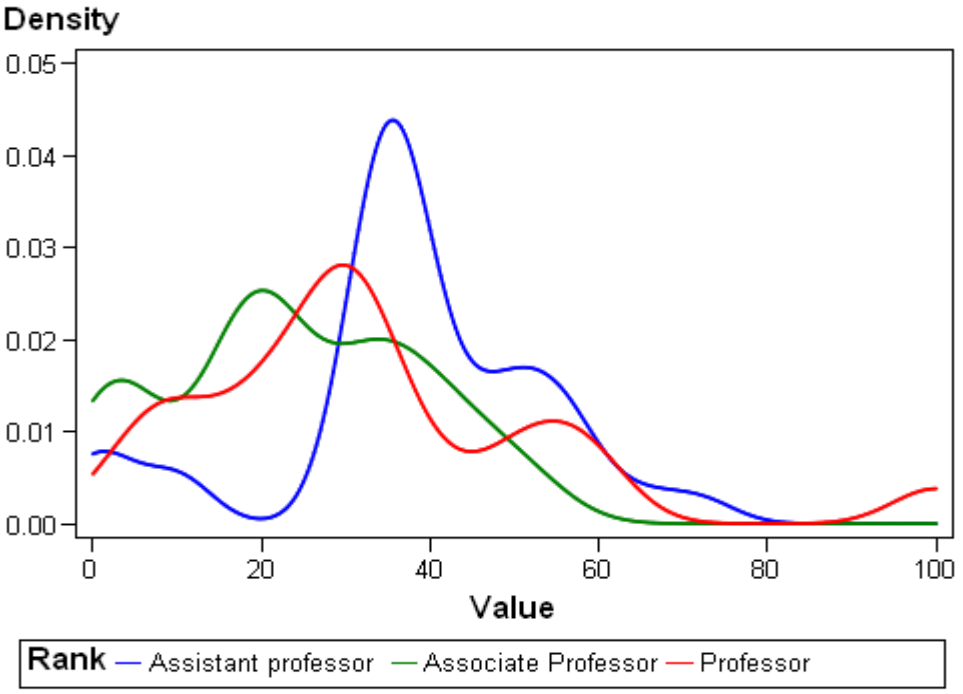


Figure 7: Research by Rank

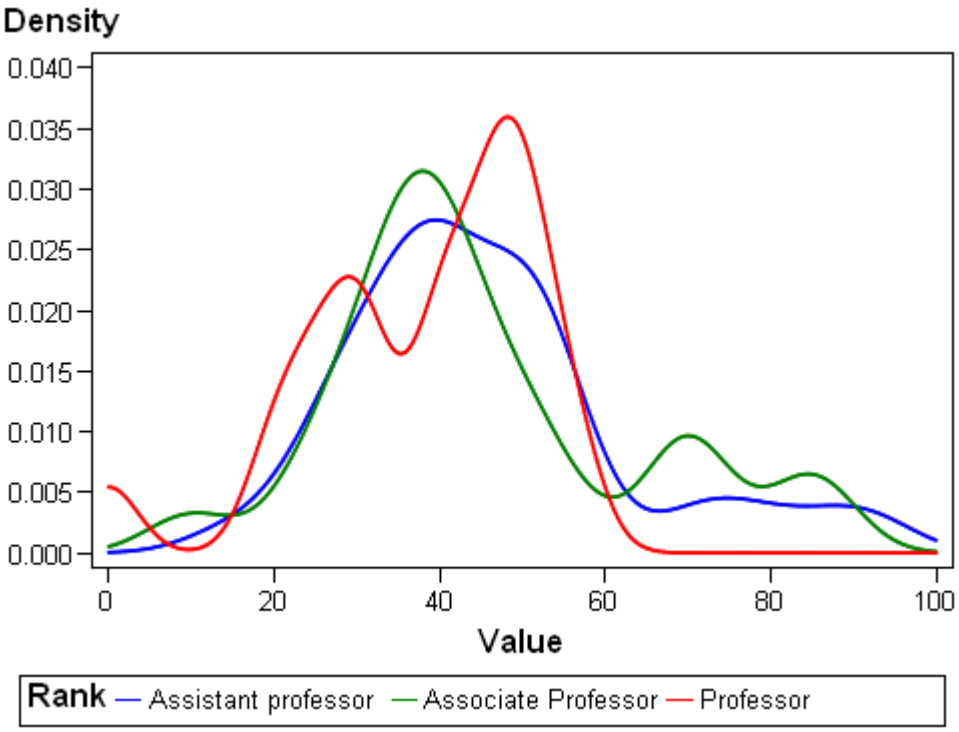


Figure 8: Teaching assignment by Rank

As figure 8 shows professor teaching density is consternated on those two peaks, while the Assistant professor illustrate more variability.

3. Data Mining

Time allocation analysis is not informative enough about the assignments nature; in this section we will investigate research, service, and administrative activities by Department.

3.1 Service

Table 5 gives a grouping of the service assignments. Most of the service responsibilities (58) focus on coordination and student advising. The next highest number (24) has to do with reviewing proposals and manuscripts. There are some who speak to recruit students (10) and those who sit on boards (13). There are remaining handfuls who are listed as giving lectures. These should be combined with those who recruit students in cluster 2. Therefore, we can reduce the six clusters to a total of four. We want to see if there is a different in cluster membership with relationship to the two Departments. Figure 9 gives the concept links related to administration.

Clusters

#	Descriptive Terms	Freq	Percentage	RMS Std.
1	coordinator, reader, fair reader, co-coordinator, coordination, symposium, webmaster, student advising, student, advising	58	0.517857142...	0.1848407...
2	fair judge, + uofl speaker, + speaker, + presentation, uofl, profession, profession committee, trip, stusent advising, stusent	10	0.089285714...	0.2464886...
3	mcgraw hill, mcgraw, hill, chapter reviewer, chapter, student advising, student, reviewer, advising, committee	3	0.026785714...	4.4554151...
4	board, chair, board member, supervisor, fair board, assistant chair, assistant, director, member, coaching	13	0.116071428...	0.1817813...
5	+ lecture, + community lecture, professional, + medium, + medium relation, + professional lecture, + relation, + seminar, consultant, community	4	0.035714285...	0.0534309...
6	proposal, nsf proposal reviewer, nsf, steering, associate, associate editor, community steering committee, proposal reviewer, + manuscript, + manuscript reviewer	24	0.214285714...	0.1774007...

Table 5: Service Assignments by Group

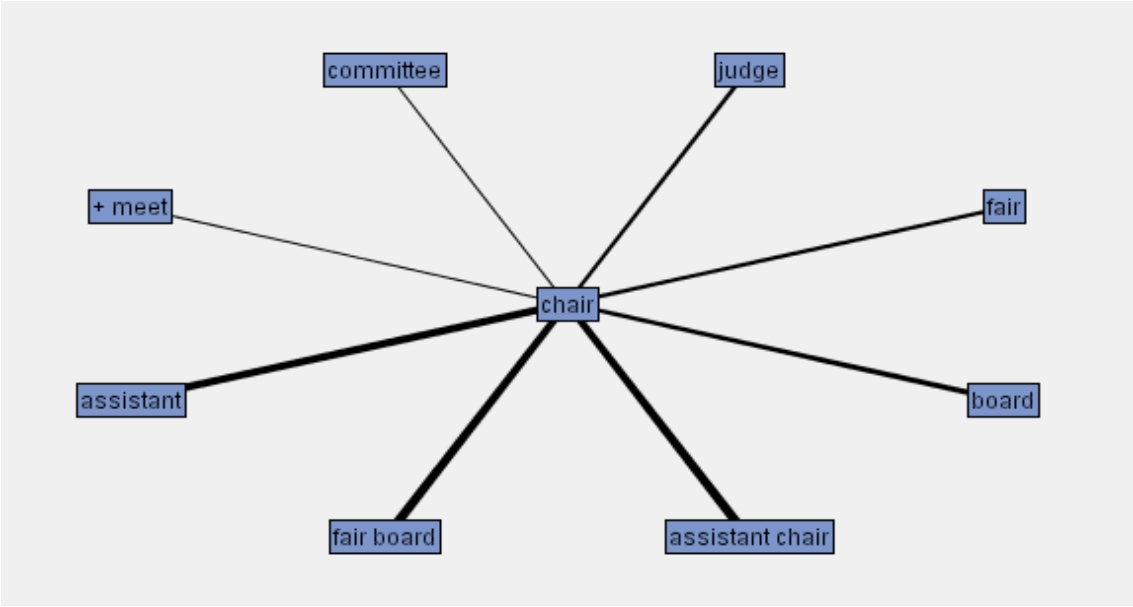


Figure 9: Concept Links for Administration

The position of Chair is related to assistant chair and to meetings. It suggests that the role of Chair is not fully spelled out. To discover the job description, we need to examine the university bylaws and personnel policies. Table 6 gives the different positions in the Department of Mathematics that include administrative time. The Department of Biology only has a position for department chair.

Administration				
Administration	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
Assistant Chair	2	18.18	2	18.18
Chair	2	18.18	4	36.36
Graduate Studies Director	2	18.18	6	54.55
Undergraduate advising coordinator	1	9.09	7	63.64
Undergraduate director	2	18.18	9	81.82
internship director	2	18.18	11	100.00

Table 6: Administrative Positions in the Department of Mathematics

At the same time, the Department of Mathematics has considerably fewer majors compared to Biology, which brings into question just why so many administrative and advising positions are required.

3.2 Research

Table 7 shows the groups of research assignments. Most (54) are defined by papers either to start, continue, or complete a project. In addition, this group includes the preparation of grant proposals. An additional 23 are directed toward the USDA (United States Department of Agriculture). While we will also see if there is a difference between departments in the definition of these groups, it is fairly likely that support from the USDA is solicited by members of the Department of Biology. Another 11 concentrate on collaborations and developing a project with a collaborator. Ten focus on a book chapter or on editing a book. There are 7 each for textbook writing and for editing.

Clusters

#	Descriptive Terms	Freq	Percentage	RMS Std.
1	research, + presentation, papaers, completing/continue project, completing/continue, nih/nsf, nih/nsf proposal, preparation, sabbatical, set up	54	0.482142857...	0.2013753...
2	textbook wirting, textbook, nsf project, wirting, continue, nsf, project, research	7	0.0625	0.2116748...
3	lead editor, lead, editor, current, keep up, previous, previous collaboration, with, workshop, + current paper	7	0.0625	0.2537610...
4	project collaboration, collaboration, develop, project, proposal, submi	11	0.098214285...	0.0927131...
5	nsf/usda, fund, internal, international, international meeting, meeting, nsf/usda proposal, nsf/usda propsal , internal fund, + manuscript, continue	23	0.205357142...	0.1832679...
6	book, chapter, + article, book chapter, book revision, prposal, wirte, revision, submit, proposal	10	0.089285714...	0.2211997...

Table 7: Research Assignments

We also look at some concept links related to research. Figure 10 examines the relationship to the term, ‘reasearch’. They include manuscript, submit, paper, and presentation. Table 8 summrizes research clusters by department which show that Mathematics tends to be very concentrated in cluster 1 while Biology has more variability in the accepted research activities.

Table of Research Clusters by Department			
Research Clusters	**Department**		
Frequency **Row Pct** **Col Pct**	**Biology**	**Mathematics**	**Total**
1	12 25.00 23.08	36 75.00 66.67	48
2	3 42.86 5.77	4 57.14 7.41	7
3	4 57.14 7.69	3 42.86 5.56	7
4	9 81.82 17.31	2 18.18 3.70	11
5	20 86.96 38.46	3 13.04 5.56	23
6	4 40.00 7.69	6 60.00 11.11	10
Total	52	54	106

Table 8: Research Clusters by Department

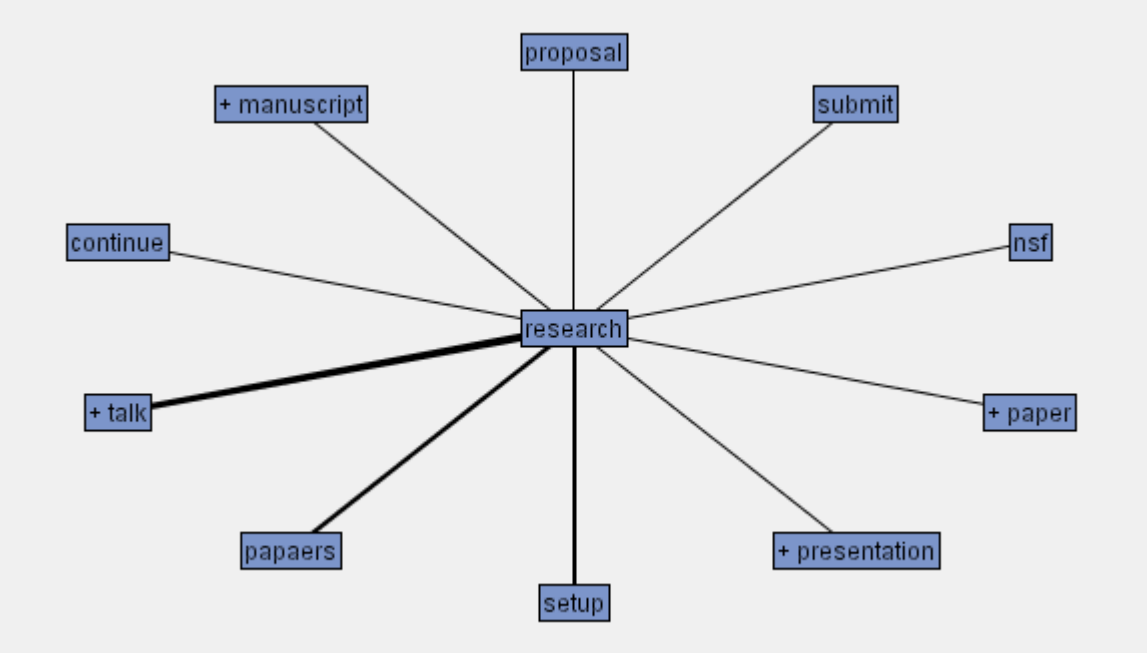


Figure 10: Concept Links to Research

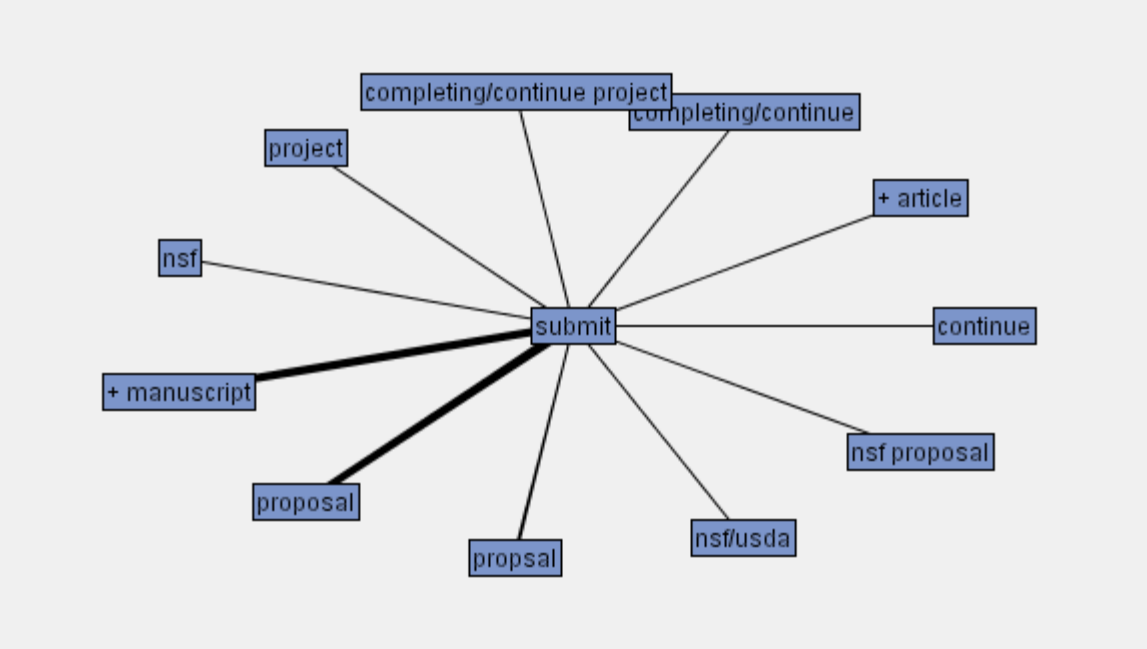


Figure11: Concept Links to Submit

The term ‘submit’ is linked to proposal, manuscript, and article. These two concept links indicate that that submission is linked to some type of manuscript or proposal. There is an expectation of publication.

3.3 Teaching

Table 9 gives a summary of the mentoring of students in the two Departments as well as the total amount of time allocated for instruction, which consists of course preparation, mentoring students, and classroom teaching.

Dep	Rank	N Obs	Variable	Mean	Std Dev	Minimum	Maximum	Sum
Bio	Assistant professor	16	Mentoring time Instruction time	0.1025 0.4425	0.0300 0.0881	0.0500 0.2700	0.1600 0.6100	1.6400 7.0800
	Associate Professor	15	Mentoring time Instruction time	0.0873 0.5486	0.0728 0.1463	0 0.3700	0.2700 0.8000	1.3100 8.2300
	Professor	17	Mentoring time Instruction time	0.0741 0.5000	0.0484 0.1193	0 0.2900	0.1400 0.6800	1.2600 8.5000
Math	Assistant professor	18	Mentoring time Instruction time	0.0405 0.5716	0.0279 0.1418	0 0.3500	0.0800 0.9300	0.7300 10.2900
	Associate Professor	7	Mentoring time Instruction time	0.0271 0.5771	0.0236 0.2146	0 0.4000	0.0600 0.8900	0.1900 4.0400
	Professor	21	Mentoring time Instruction time	0.0314 0.3938	0.0519 0.1604	0 0	0.2000 0.5700	0.6600 8.2700

Table 9: Summary of Teaching

It shows that Biology spends considerably more time in one-on-one mentoring compared to Mathematics. One of the reasons is that Biology has many more graduate students who need mentoring. Biology gives assistant professors a lighter instructional load; associate professors have a higher load. In contrast, professors in Mathematics have a considerably lighter instructional load, partially because administrative positions are accompanied by course release, and the course release creates outliers. Figure 12 gives a kernel density estimation of the overall instructional load. It indicates that the distributions of both departments are similar.

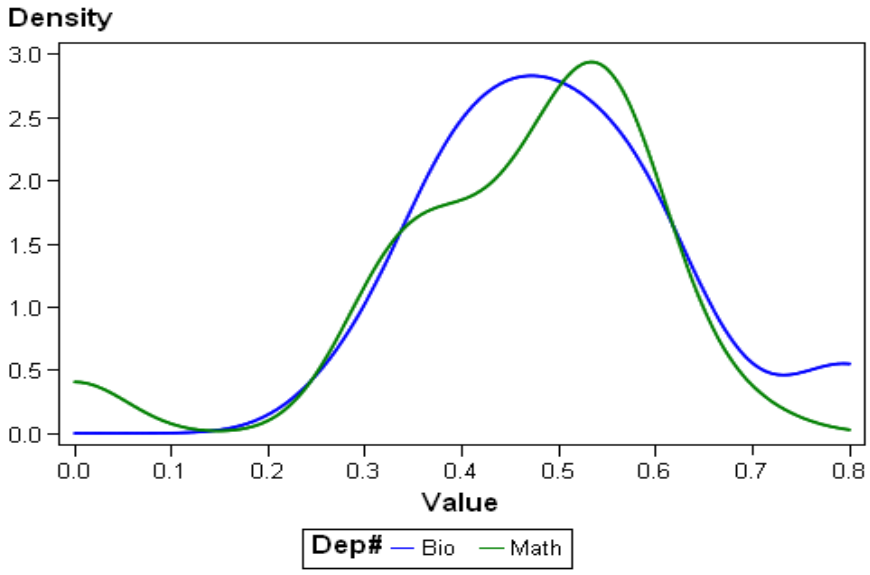


Figure 12: Instructional Time by Department

Figures 13-15 give the distributions of instruction by rank. Mathematics has more variability in the teaching assignments compared to Biology for associate professors. Figure 15 for full professor shows that Mathematics has a definite shift to lower teaching effort; both departments show a bimodal distribution. In Mathematics, the lower instructional load is related to administrative positions; in Biology, it is related to research grants.

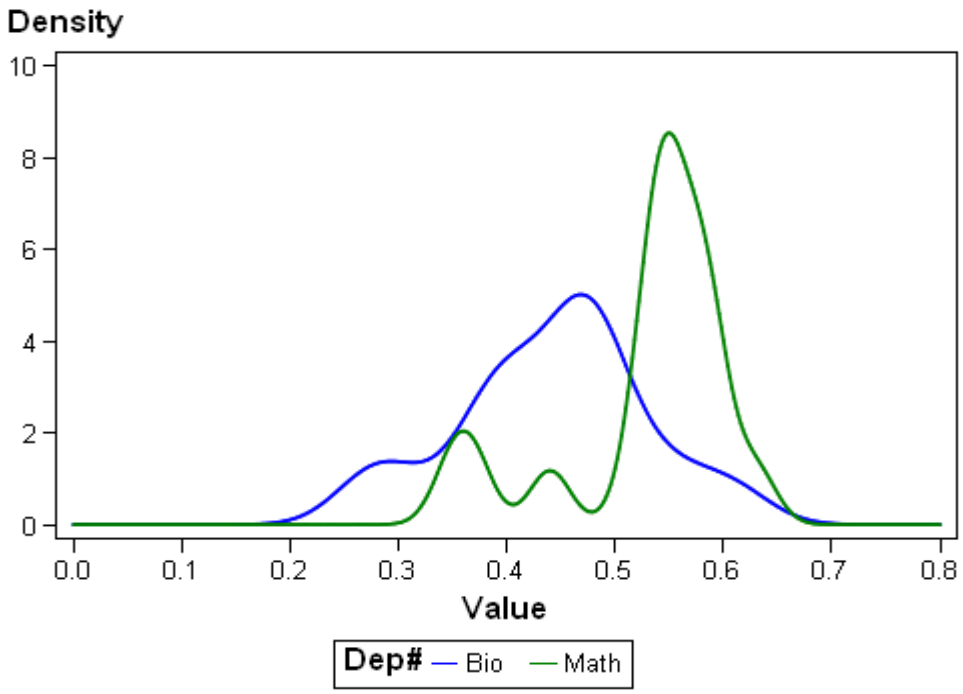


Figure 13: Instructional Time for Assistant Professors

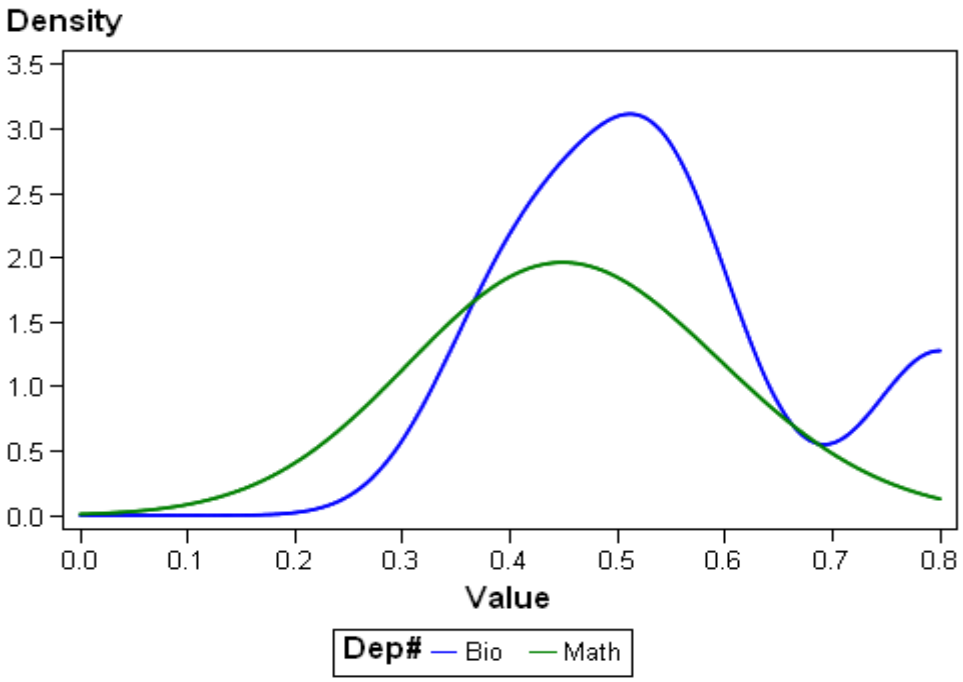


Figure 14: Instructional Time for Associate Professors

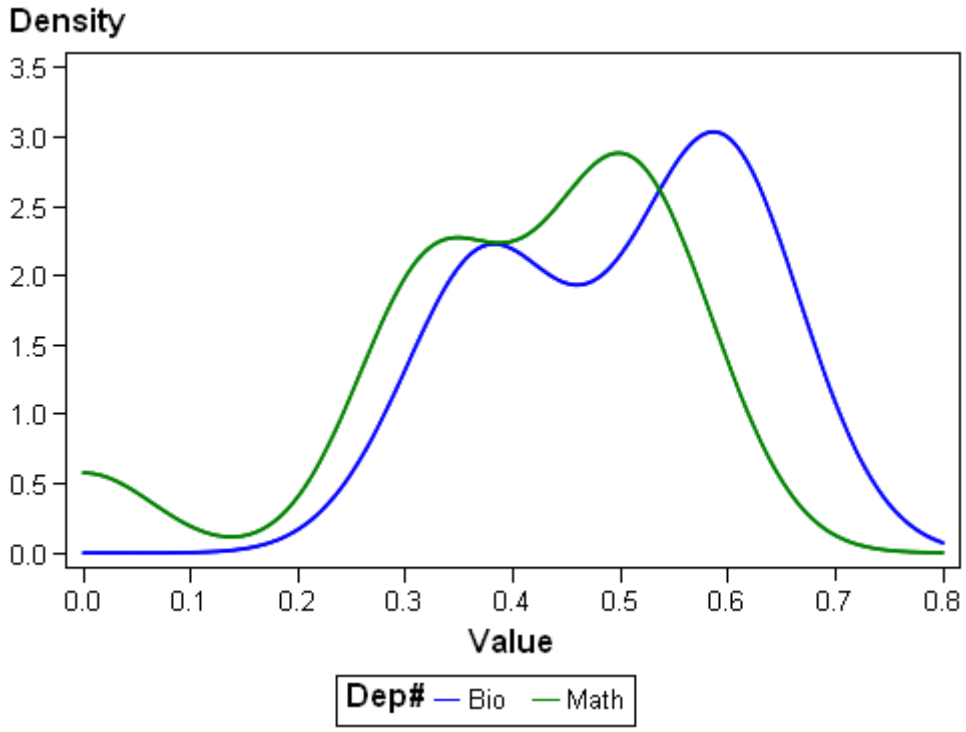


Figure 15: Instructional Time for Full Professors

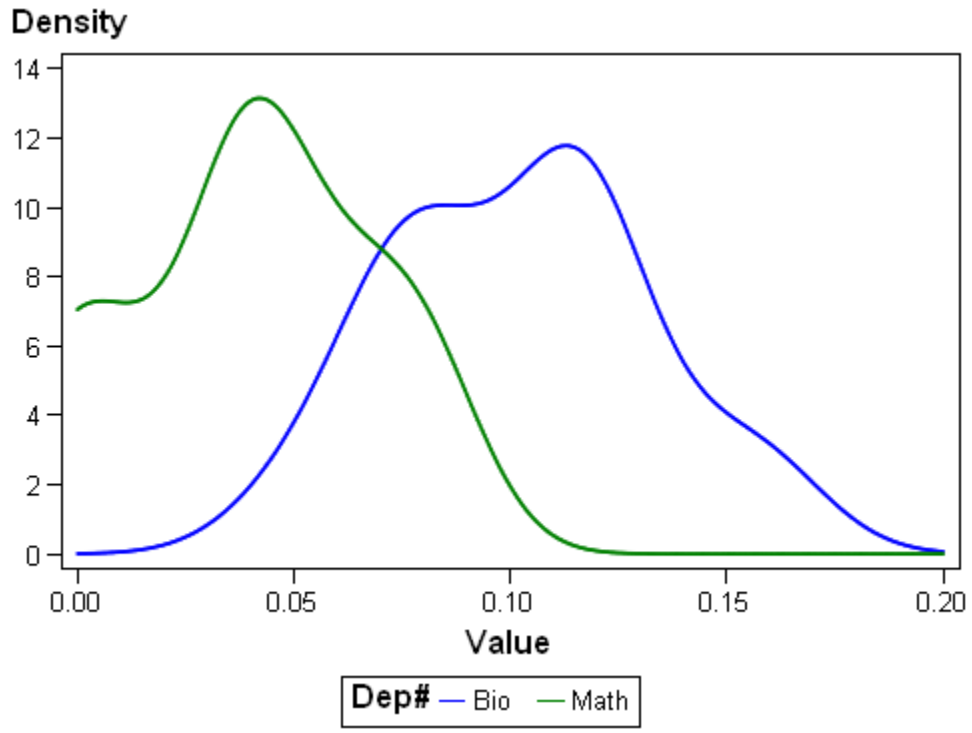


Figure 16: Mentoring Time for Assistant Professors

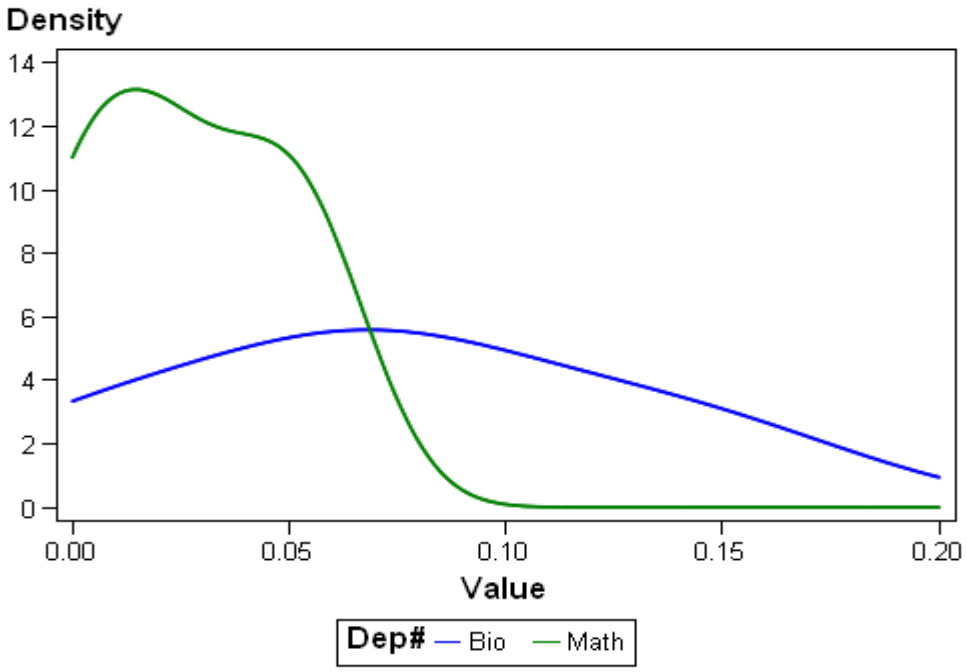


Figure 17: Mentoring Time for Associate Professors

Figures 16-18 show the level of mentoring by rank in the two departments. Assistant professors in Biology spend considerably more time mentoring compared to faculty in

Mathematics. Biology clearly has more variability in the level of mentoring, and can have almost 20% time for mentoring. Again, Biology spends considerably more time mentoring students.

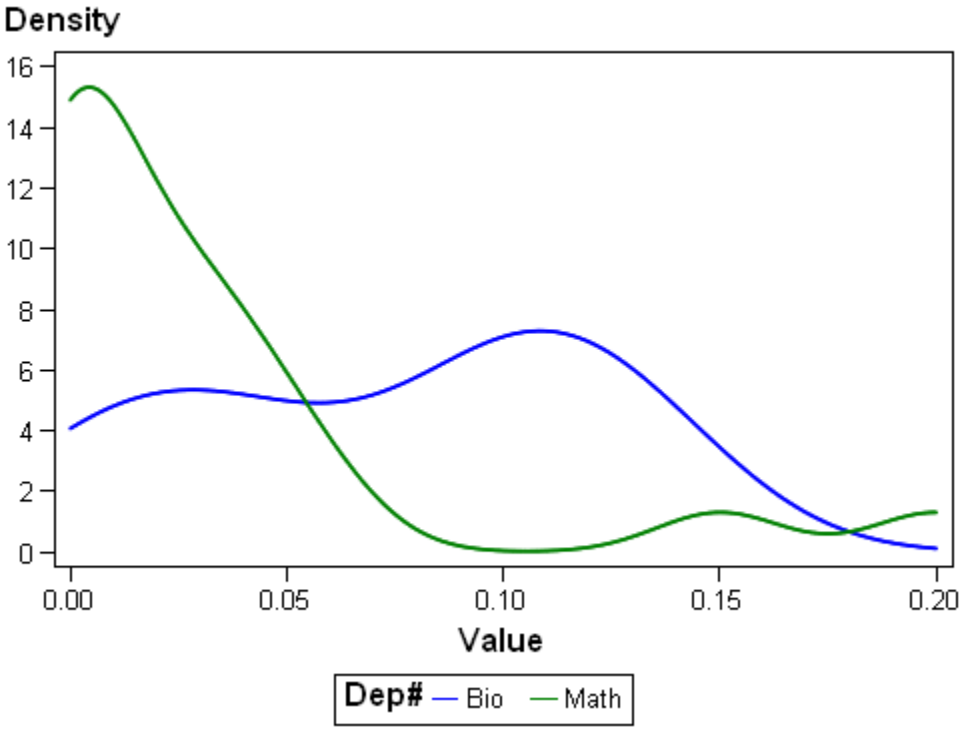


Figure 18: Mentoring Time for Professors

Clusters

#	Descriptive Terms	Freq	Percentage	RMS Std.
1	+ 205, + 640, + 301	13	0.116071428...	0.1195724...
2	+ 102-01, + 501, + 301	20	0.178571428...	0.0844955...
3	+ 301, + 501	15	0.133928571...	0.0680357...
4	+ 640, + 501, + 301	17	0.151785714...	0.0859013...
5	+ 501, + 640	47	0.419642857...	0.0549268...

Table 10: Teaching Clusters by Course Levels

Table 10 gives the teaching clusters defined in terms of course level. It shows that there are three clusters that include graduate courses (600-level) and two clusters that do not. We want to see if there are differences by rank and by Department in the assignment of such courses (Table 7).

Table of Teaching Clusters by Department			
Teaching Clusters	**Department**		
Frequency **Row Pct** **Col Pct**	**Biology**	**Mathematics**	**Total**
1	9 69.23 17.31	4 30.77 7.41	13
2	3 15.00 5.77	17 85.00 31.48	20
3	12 80.00 23.08	3 20.00 5.56	15
4	17 100.00 32.69	0 0.00 0.00	17
5	11 26.83 21.15	30 73.17 55.56	41
Total	52	54	106

Table 11: Teaching Cluster by Department

Table 11 shows that Mathematics is largely concentrated in two of the 5 clusters; clusters 2 and 5. One contains graduate courses, the other does not. It suggests that there are faculty who are assigned graduate courses regularly while the other faculty a not assigned 600-level courses. We want to examine this in more detail.

Table of Teaching Cluster by Rank				
Teaching Cluster	**Rank(Rank)**			**Total**
Frequency **Row Pct** **Col Pct**	**Assistant professor**	**Associate Professor**	**Professor**	
1	1 10.00 2.85	8 80.00 36.36	1 10.00 2.63	10
2	8 44.44 22.22	5 27.78 22.73	5 27.78 13.16	18
3	4 26.67 11.11	4 26.67 18.18	7 46.67 18.42	15
4	8 50.00 22.22	2 12.50 9.09	6 37.50 15.79	16
5	15 40.54 41.67	3 8.11 13.64	19 51.35 50.00	37
Total	36	22	38	96

Table 12: Teaching Clusters by Rank

Full professor tend to concentrate in cluster 5 with graduate courses; assistant professor are divided into clusters 2 and 4; one cluster with graduate courses and the other without. We also look at the relationship of teaching cluster to research cluster (Table 13).

Teaching Clusters	Research Clusters						
Frequency Row Pct Col Pct	1	2	3	4	5	6	Total
1	6 46.15 11.11	1 7.69 14.29	1 7.69 14.29	1 7.69 9.09	3 23.08 13.04	1 7.69 10.00	13
2	14 70.00 25.93	1 5.00 14.29	3 15.00 42.86	1 5.00 9.09	1 5.00 4.35	0 0.00 0.00	20
3	2 13.33 3.70	0 0.00 0.00	1 6.67 14.29	3 20.00 27.27	6 40.00 26.09	3 20.00 30.00	15
4	2 11.76 3.70	2 11.76 28.57	2 11.76 28.57	1 5.88 9.09	9 52.94 39.13	1 5.88 10.00	17
5	30 63.83 55.56	3 6.38 42.86	0 0.00 0.00	5 10.64 45.45	4 8.51 17.39	5 10.64 50.00	47
Total	54	7	7	11	23	10	112

Table 13: Teaching Clusters by Research Clusters

Table 13 shows that most of the faculty are in research cluster 1 and most of these faculty are in teaching clusters 1, 2, and 5. However, faculty in research cluster 5 is more likely to be in teaching clusters 3 and 4. These faculty are all in the Department of Biology. These two research clusters are related to collaboration and editing. These same faculty teach advanced, graduate courses.

4. Research Productivity

We also want to examine research productivity by examining faculty curriculum vita.

Rank	Dept	N Obs	Variable	Mean	Std Dev	Minimum	Maximum	N
Assistant professor	Bio	20	IndpSsupervison_pre	2.1000	2.1496	0	6.0000	20
			# students supervised	1.9500	2.2354	0	6.0000	20
			Thesis_S_NO	2.2500	1.6503	0	5.0000	20
			Thesis_pre	3.7500	2.8631	0	9.0000	20
	Math	18	IndpSsupervison_pre	0.6111	1.1447	0	4.0000	18
			# students supervised	0.2941	0.5878	0	2.0000	17
			Thesis_S_NO	0.1111	0.3233	0	1.0000	18
			Thesis_pre	0.2222	0.6467	0	2.0000	18
Associate Professor	Bio	15	IndpSsupervison_pre	0.7333	0.7988	0	2.0000	15
			# students supervised	1.0666	1.1629	0	3.0000	15
			Thesis_S_NO	4.0666	4.2167	0	14.0000	15
			Thesis_pre	5.2000	5.2399	0	20.0000	15
	Math	9	IndpSsupervison_pre	1.3333	4.0000	0	12.0000	9
			# students supervised	0.6666	2.0000	0	6.0000	9
			Thesis_S_NO	0.6666	1.3228	0	3.0000	9
			Thesis_pre	1.2222	2.4381	0	6.0000	9
Professor	Bio	17	IndpSsupervison_pre	1.3529	1.4116	0	4.0000	17
			# students supervised	1.5882	1.6224	0	4.0000	17
			Thesis_S_NO	2.7941	2.6164	0	8.0000	17
			Thesis_pre	3.9411	4.1150	0	12.0000	17
	Math	25	IndpSsupervison_pre	0.5000	2.0641	0	10.0000	24
			# students supervised	0.2916	1.0826	0	5.0000	24
			Thesis_S_NO	0.5416	1.1412	0	5.0000	24
			Thesis_pre	1.2083	2.4491	0	10.0000	24

Table 14: Research activity related summary

As it shown on table 14, the research related activities of biology department shows higher means across the different faculty rank.

Dept: Biology

Variable	Mean	Std Dev	Minimum	Maximum	N
Teaching percentage	40.4423	15.3065	10.0000	80.0000	52
Research percentage	34.4038	18.8744	1.0000	72.0000	52
Services percentage	11.7788	6.5674	1.0000	28.0000	52

Dept= Mathematics

Variable	Mean	Std Dev	Minimum	Maximum	N
Teaching percentage	44.6274	17.8168	0	93.000	51
Research percentage	30.8627	19.2145	0	100.000	51
Services percentage	13.5400	9.5622	0	50.000	50

Table 15: Teaching, Research, and services activity summary

Table 15 shows the Teaching, Research and Service percentage on department level; it is clear that for the Biology department the teaching percentage in average less than the teaching percentage in Mathematics department, while is the situation is reserved on

research aspect. Which highlight the importance of research on Biology department and availability of funds. It also shows that the services roles on the Biology department are less then the Mathematics department.
The dispersion for the Biology department is less then dispersion for the Mathematics department which reflect unbalanced faculty work load. As we can see the service percentage for Mathematics department can is double the service percentage for the Biology department.

Dept: Biology

Analysis Variable : Publication numbers						
Rank	**N Obs**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**N**
Assistant Professor	20	2.2000	4.0600	0	14.0000	20
Associate Professor	15	2.7333	5.4177	0	15.0000	15
Professor	17	4.2941	7.6873	0	24.0000	17

Dept= Mathematics

Analysis Variable : Publication numbers						
Rank	**N Obs**	**Mean**	**Std Dev**	**Minimum**	**Maximum**	**N**
Assistant Professor	18	2.8888	3.6118	0	12.0000	18
Associate Professor	9	1.7777	3.6666	0	10.0000	9
Professor	25	7.2400	12.4675	0	46.0000	25

Table 16: Publication number summary

Number of publication is a good representation for research productivity, the above table shows interesting results; in the case of Assistant Professor we can see that although that the maximum value of Biology department is larger then the correspondence value for Mathematics department while the mean value for Assistant Professor/ Mathematics department is lower than the mean value for Assistant Professor/ Biology department which show less variability. In the case of Associate Professor the average of Associate Professor/Biology shows more variability which can be due to the lower number of Associate Professor for Mathematics department. The same observation is valid for the Professor rank/Mathematics department.

Teaching is the main concern for any educational organization, number of graduating student is an excellent illustration for department productivity. Table 17 shows number of student graduating on different programs. As it demonstrates total number of student at Biology department are more than double the number for Mathematics department, also it is clear that the BS & BA student -which require more teaching load- for Biology department are three times the number for Mathematics department see Figure 18.

	BA	**BS**	**MA**	**MS**	**PhD**	**Total**
BIO	30	30	0	2	1	63

Math	7	15	4	0	1	27

Table 17: Graduating student number summary

In general the relation between student number and administration positions are escalating, yet Mathematics department not compelling with that rule. As shown on Table six Biology department administrations positions are much less what mathematics department has.

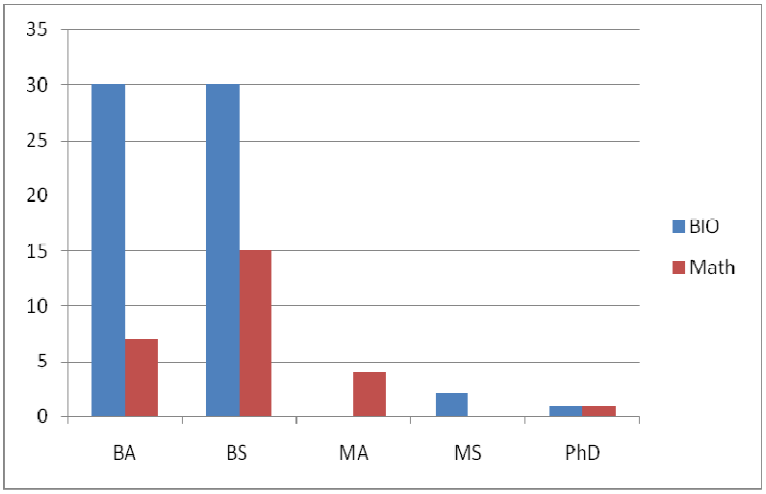


Figure 18: Graduating student number per department

Many organizations uses external fund number as a metrics for faculty and schools productivity. The results show how Biology department stresses external funding while mathematics department has less stresses for external funding. Table 18 illustrates that 55% of the mathematics department faculties have no external funding and didn't submit any.

Grants				
Grant	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
0	16	55.17	16	55.17
1	10	34.48	26	89.66
2	2	6.90	28	96.55
3	1	3.45	29	100.00

Table 18: Mathematics department external fund summary

Discussion

It is clear that there are differences in faculty workload assignments between Biology and Mathematics. While Biology stresses external funding, Mathematics does not. Instead, Mathematics has defined a need to have more administrative positions that serve to reduce the overall teaching workloads for senior faculty members in the Department. These administrative positions bring a reduction in the research productivity as well. Because grants are not stressed in Mathematics in the same way that is in Biology, Biology has considerably more external funding.

As it shown number of student supervised or enrolled on graduate program for Mathematics department are lower thank Biology department which is due to less number of external funding as it also considered as good source to enrich department facilities.

Annual work plans (AWP) and Curriculum Vitae are the main sources of information for faculty workload study and analysis. It reveled that it is very important to design AWP in such away to be more informative and uniform epically for research sections. Also it is vital to update and improve faculty's feedback mechanisms and policy to ensure research and teaching quality assurance.

Internet Investigations of Text Material to Compare Programs Across Institutions

Patricia B. Cerrito
Department of Mathematics
University of Louisville
Louisville, KY 40292
502-852-6010
502-852-7132 (fax)
pcerrito@louisville.edu

Abstract
Institutions and their faculty now post considerable amounts of information on their web sites. Unfortunately, most of it is difficult to extract and to use to make comparisons across institutions. Web and text mining tools have been developed that can be used to investigate this information, and to make institutional comparisons. We demonstrate the feasibility of these tools by using them to examine course syllabi in Mathematics. We find that the syllabi remain very traditional and depend largely on quizzes and examinations for assessments of student progress. We also show that mathematics degree programs remain very traditional, examination-based rather than research focused.

Introduction

We use web and text mining tools to examine institutional expectations concerning course syllabi and degree programs to demonstrate how such automated tools can be used to compare institutions. We focus our search on mathematics, and compare the results to other disciplines. The information we are seeking is readily available on university websites. However, it is difficult and time-consuming to extract, as much of it is contained within text documents with non-standardized means of locating the information within each individual web site.

While only a small number of publications have discussed the content of a mathematics PhD program and just a handful have discussed the content of a mathematics syllabus, there are few faculty in the discipline who incorporate any suggested changes. We show that mathematics instruction remains very traditional, although some departments are moving to a more research oriented approach. This is quite contrary to the considerable changes that have occurred in K-12 mathematics education with the National Council of Teachers of Mathematics' standards of mathematics education.

Method

The first step in the process of investigation is to perform a keyword search on the term, "mathematics syllabus and university". Because of the preponderance of advertisements that now appear at the beginning of a returned search, we need to include enough search terms to minimize the appearance of ads. Then, the SAS macro, %tmfilter, is used to collect the returned documents into a directory,

```
%tmfilter(url=http://www.ask.com/web?q=mathematics+syl
labus+and+university&sm=adv&qsrc=196,
    depth=1,
    dir=c:\syllabi\dir,
    destdir=c:\syllabi\destdir,
    norestrict=1,
    dataset=syllabi.documents);
```

Once the documents are placed in the directory (c:\syllabi), the document locations are stored in a SAS dataset, syllabi.documents. These locations are used to point to the directory so as to conserve memory storage. We did a second search to investigate mathematics PhD degree requirements using the search, http://www.ask.com/web?q=mathematics+phd+degree+requirements&search=search&qsrc=0&o=0&l=dir.

Once the web sites have been gathered into a document folder, they are analyzed using the process of text parsing and singular value decomposition. The purpose of these two steps is to remove terms from the documents that do not contribute to any meaningful analysis, terms such as "of", "and", "the". Parsing then involves creating a term by document matrix. As this matrix is extremely large, singular value decomposition

compresses the matrix to size, say 100 (dimensions). This compression can be done since the term by document matrix is very sparce (containing mostly zeros).

Text mining uses both grammar and syntax to find documents that are similar. Different weights are given to different words (with the most commonly used terms given the least weights), and the weighted terms are used to define the documents into mutually exclusive clusters.

Results

Course Syllabi

For the course syllabi, almost 300 documents were returned. Documents from the ads at the beginning of the returned search were filtered out of the analysis. What documents remained were defined into a total of 6 groups. The term “exam” appeared 1282 times in 161 documents, and the similar term “quizzes” appeared 353 times in 102 documents and “test” appeared 911 times in 146 documents. The term “grade” appeared 1154 times in 174 documents. In contrast, “homework” appeared 698 times in 136 documents, indicating that many of the syllabi did not discuss homework, although the term “assignment” appeared 503 times in 135 documents. Figure 1 gives the clusters of documents of the mathematics syllabi.

Figure 1 shows which groups of documents focus on which items. For example, cluster 5 with 69 documents is very focused on exams and homework, as is cluster 6 with 31 documents. In contrast, cluster 1 is more focused on course content, discussing concepts such as polynomials and distributions. Documents in cluster 3 spend time on accommodations for students with disabilities.

Figure 1. Clusters of Documents

Clusters

#	Descriptive Terms	Freq	Percentage
1	+ square, + distribution, + property, + variable, + equation, + design, + polynomial, linear, + apply, + computer, basic, + system, + concept, + model, + function, + credit, + number, + method, + graph, + description	31	0.113138686...
2	+ appear, + drop, + withdraw, + receive, + college, + find, + need, + question, + date, + write, + make, + have, + material, + test, + include, + do, + objective, + other, + assignment, + graph	34	0.124087591...
3	+ last, + contact, + revise, technological, + topic, + university, tennessee technological university, + disability require accommodation, + reference, + center, first, + request, + accommodation, + possible text, roaden university center, department, + end, + locate, catalog, + prerequisite	57	0.208029197...
4	+ graph, + provide, + write, + date, + learn, + solve, + policy, + syllabus, + function, + make, + grade, + topic, + other, + have, + number, + calculator, + assignment, + hour, + final, + time	52	0.189781021...
5	+ exam, + schedule, + grade, + final exam, + final, + miss, + homework, + not, + material, + class, + give, + will, + section, + assignment, + hour, all, + day, + syllabus, + do, on	69	0.251824817...
6	+ meet, + need, + review, + email, + question, + policy, + follow, + assignment, + day, + do, + time, + have, + provide, + work, + exam, + grade, + study, + syllabus, + hour, + complete	31	0.113138686...

In addition to clusters, we use concept links to examine connections between terms used in the course syllabi. Figure 2 shows the connections to ‘final exam’. Figure 3 gives the connections to ‘several quizzes’. Figure 2 indicates a connection to the usage of calculators (‘hand held’) but no connection to computers. Figure 3 shows connections to ‘management’, ‘timely’, and ‘mastery’, indicating that quizzes are used, even in graduate courses, to instill discipline in the students.

Figure 2. Links to 'final exam'

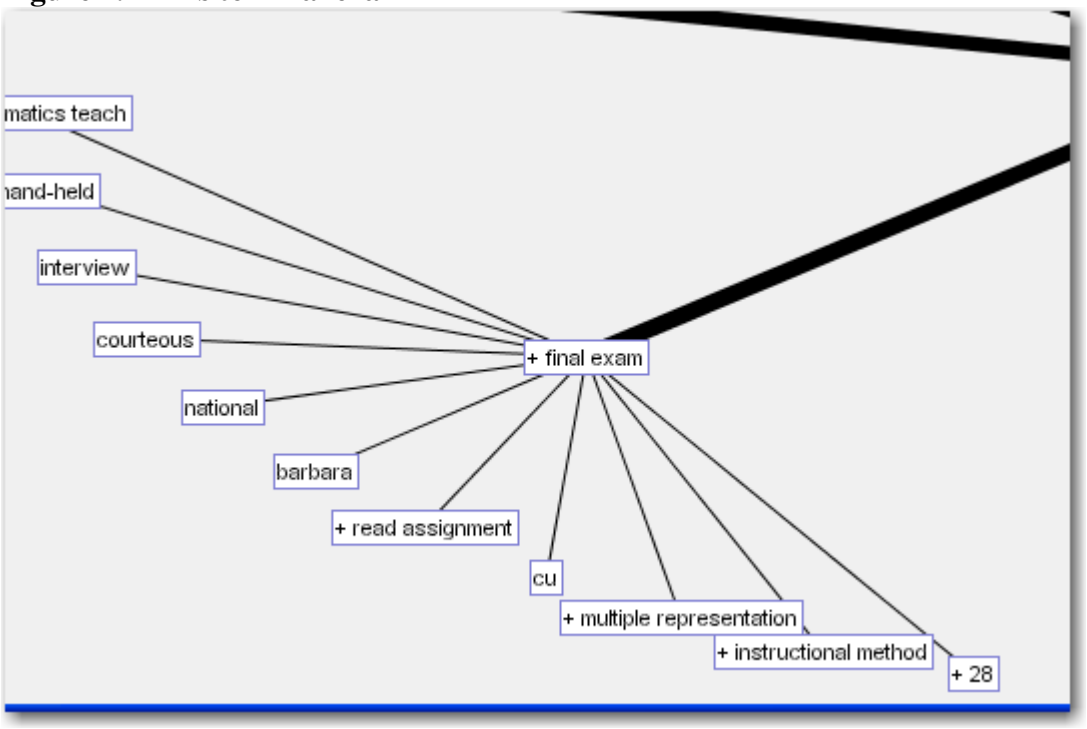
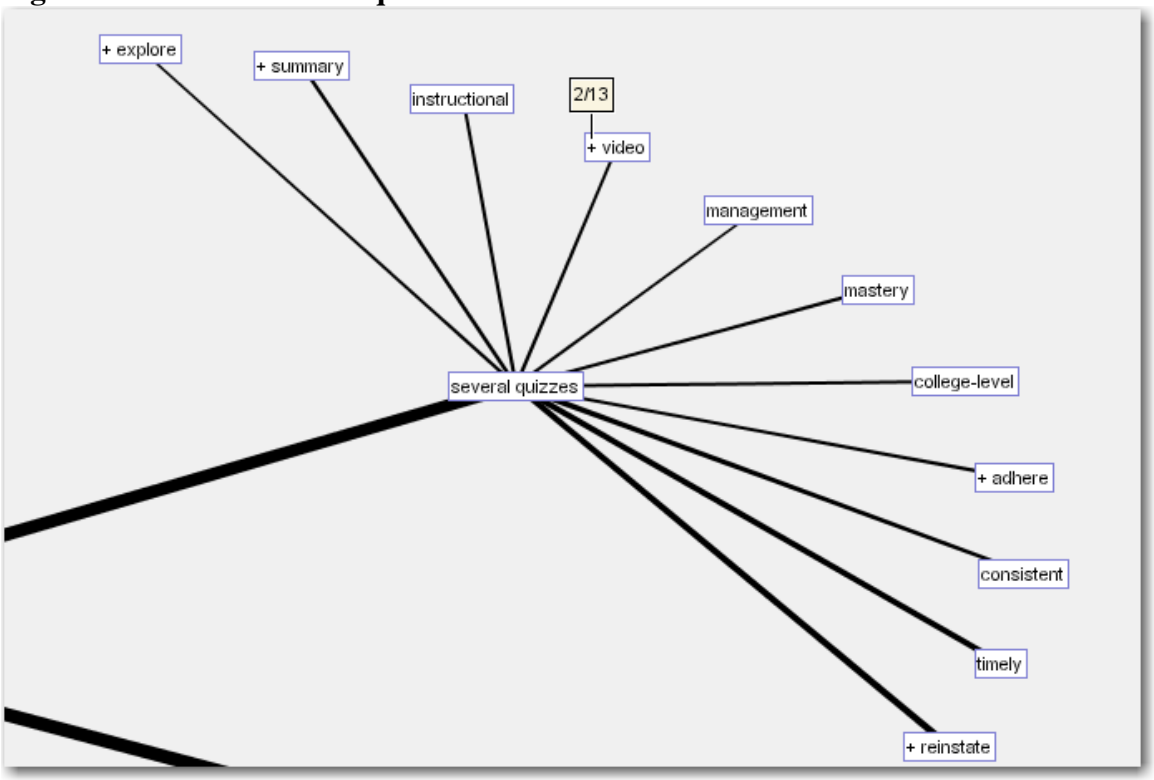


Figure 3. Links to 'several quizzes'



Similarly, Figure 4 shows connections to 'grade' and Figure 5 shows links to 'classroom activity'.

Figure 4. Links to 'grade'

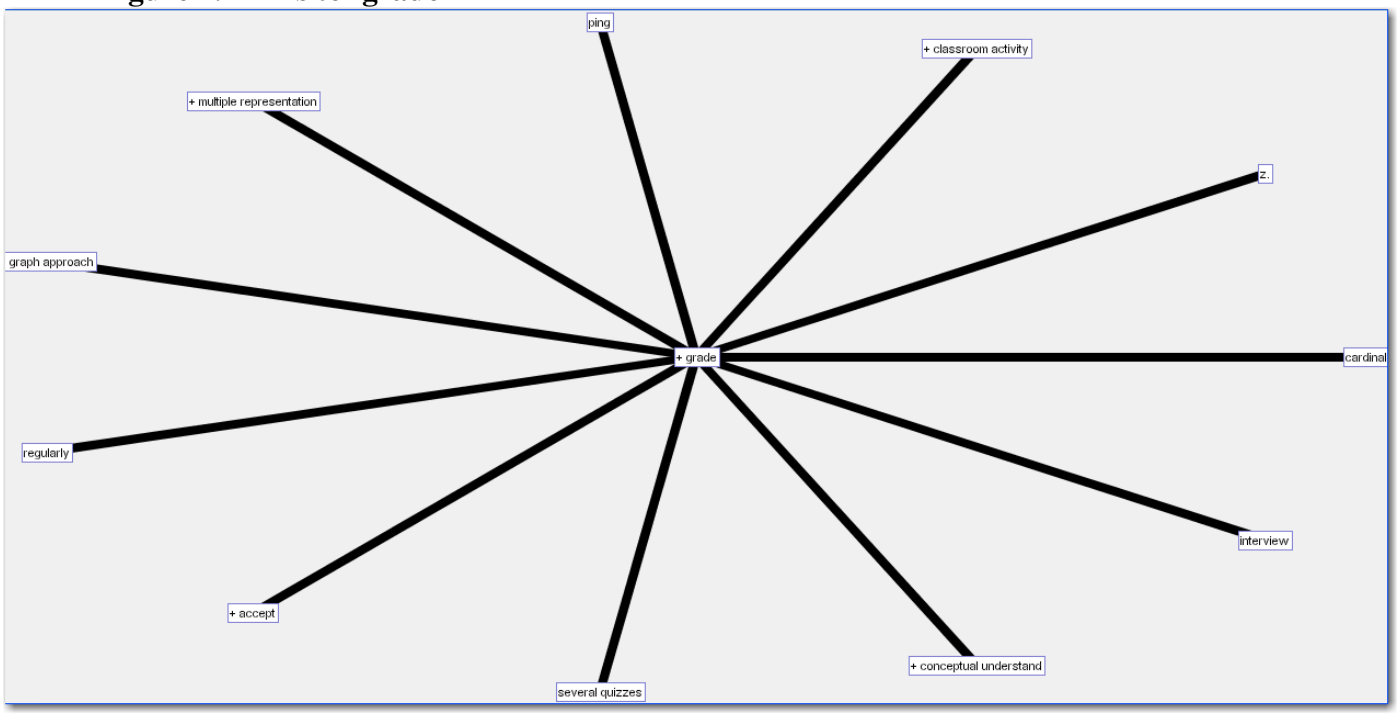
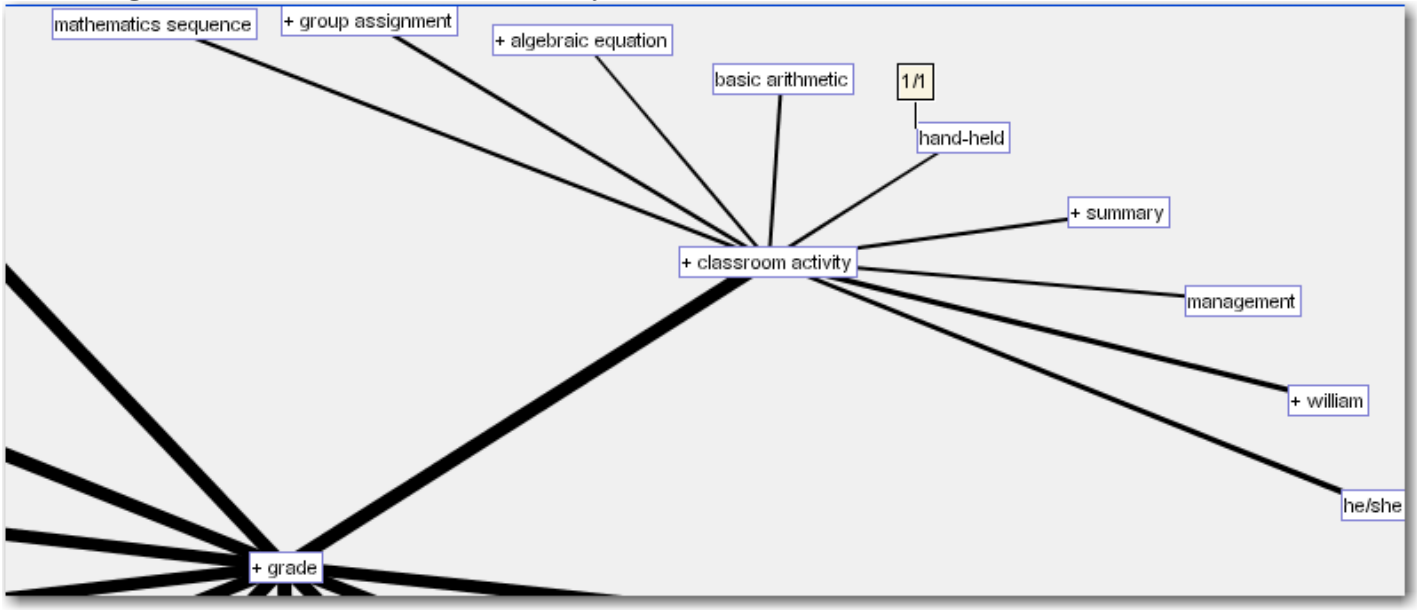
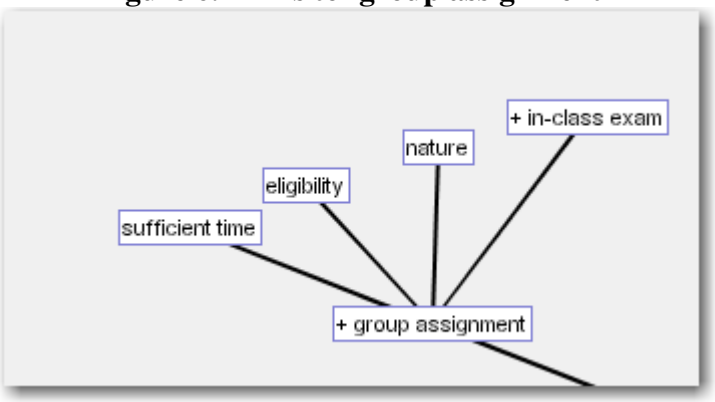


Figure 5. Links to 'classroom activity'



Grades are linked to quizzes and exams. They are also linked to 'classroom activity' and to 'regularly' (homework). 'Classroom activity' is related to 'group assignment' as well as to 'hand-held' (calculators). Therefore, there are some (just a few) course syllabi that discuss group assignments. To get a further understanding, Figure 6 shows the links to 'group assignment'. These are related to 'sufficient time' and to 'in-class exam,' indicating that the group assignment focuses on a group, in-class exam.

Figure 6. Links to 'group assignment'



PhD Requirements

When we examined the terms related to PhD degree requirements, it was found that all but two programs still relied on the model that required qualifying examinations prior to initiating a program of research. A total of 100 documents were examined, with 78 related to PhD degree requirements. The term, "examination," was listed 916 times in the documents. The term, "research" appeared only 454 times. Concept links are used to examine the relationship between examinations and research more closely.

Figure 7 gives the clusters for degree requirements. Cluster 1 with 42 documents concentrates on the examinations and other requirements.

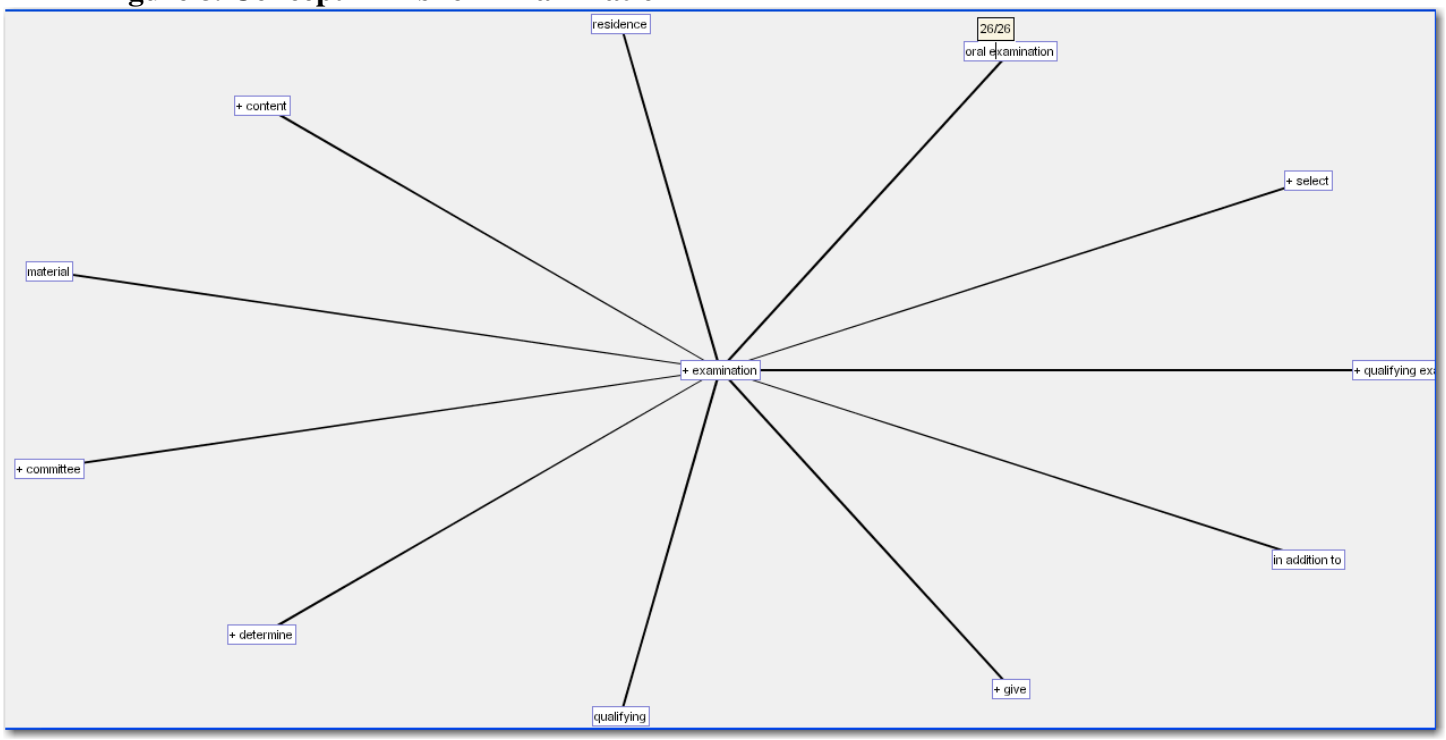
Figure 7. Clusters of Degree Requirements

Clusters

#	Descriptive Terms	Freq	Percentage
1	+ examination, + department, + master, + language, + student, + requirement, + degree, + credit, + 's, + hour	42	0.531645569...
2	+ select, + sequence, + make, + write, + expect, + member, + require, + satisfy, + master, + area	9	0.113924050...
3	+ meet, + follow, + field, + make, + apply, + science, + topic, + application, + write, + have	16	0.202531645...
4	+ hold, + professor, + component, + enroll, + part, + form, + make, + meet, + prepare, + candidate	6	0.075949367...
5	+ contact, + allow, + offer, + include, + prepare, + approve, + apply, + complete, + department, + area	6	0.075949367...

Figure 8 shows the concept links for "examination"; Figure 9 gives the links for "qualifying".

Figure 8. Concept Links for ‘Examination’



In most cases, ‘examination’ is linked to ‘qualifying’ indicating that more programs still use examinations rather than research to move students into degree candidacy. The qualifying examinations still focus on mathematics foundations (ie, pure mathematics) with a handful that also include applied mathematics. Figure 10 gives the connections to ‘dissertation’.

Figure 9. Concept Links to 'Qualify'

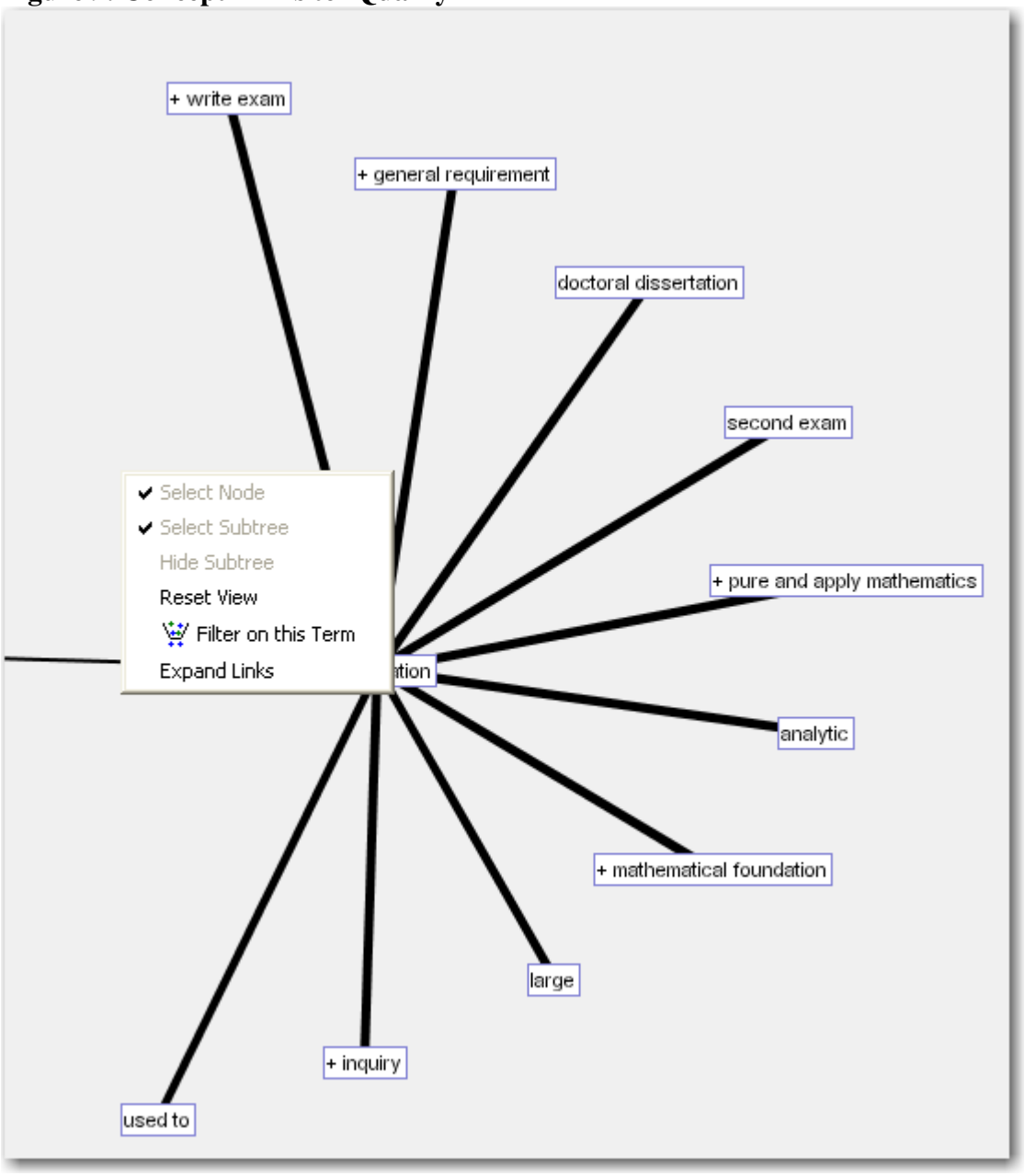
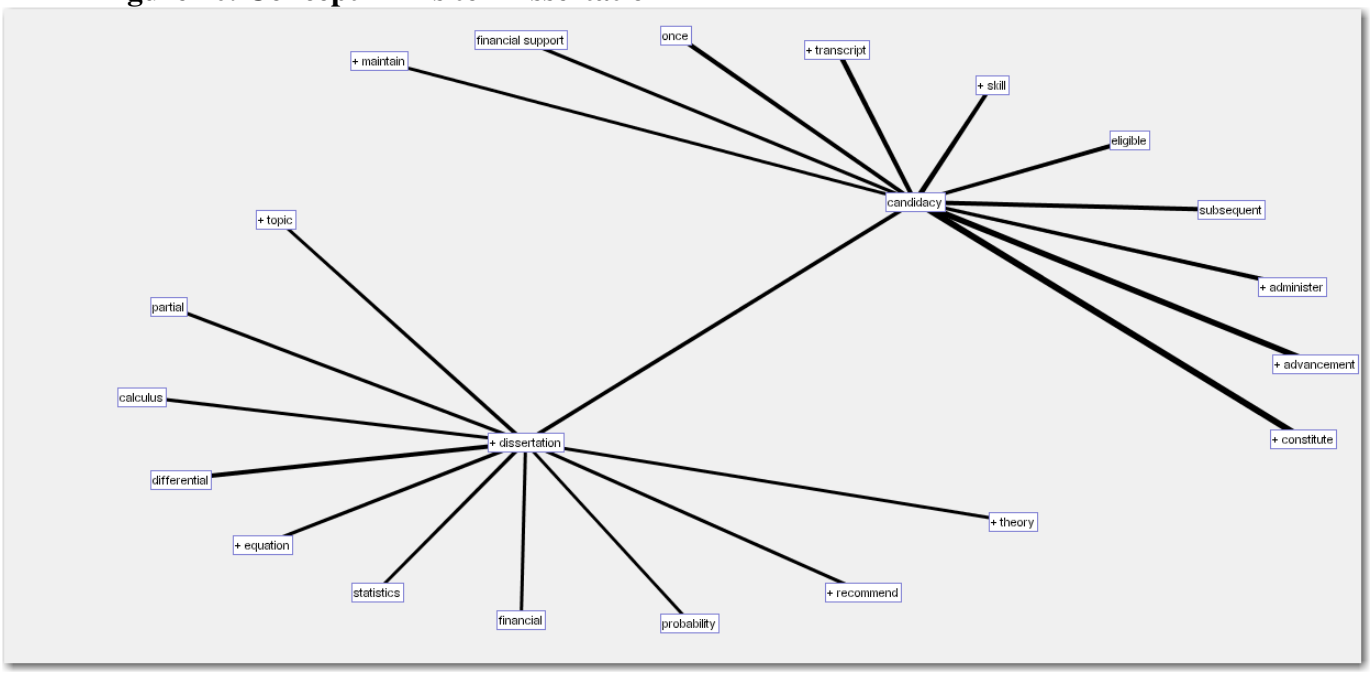
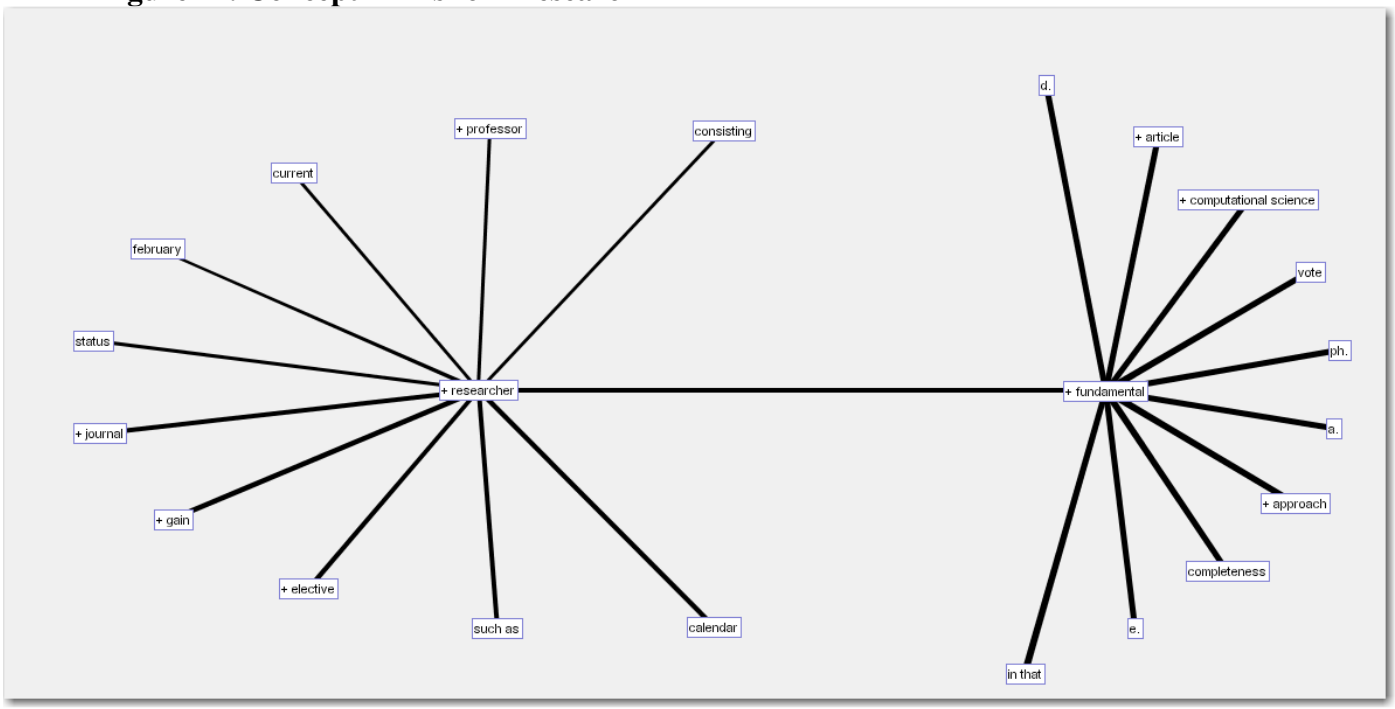


Figure 10. Concept Links to 'Dissertation'



While most of the links are to the area of mathematics, the remaining connection is to candidacy. In other words, the dissertation leads to candidacy. Figure 11 gives the connections to 'research'.

Figure 11. Concept Links for 'Research'



Interestingly**,** there is no link between research and examination. In other words, degree requirements do not make this connection.

Conclusion

The web and text mining techniques greatly simplify the ability of researchers to examine source documents so that programs and requirements can be compared across institutions in a quantitative fashion. Text mining also allows us to drill down into document details for close examination.

Since there is little to connect research to examinations (although qualifying examinations do link to dissertation), we recommend that advancement to candidacy shift to emphasize the most important aspects of a PhD degree; that is, the development of research skills rather than testing skills. Few programs have made this shift.