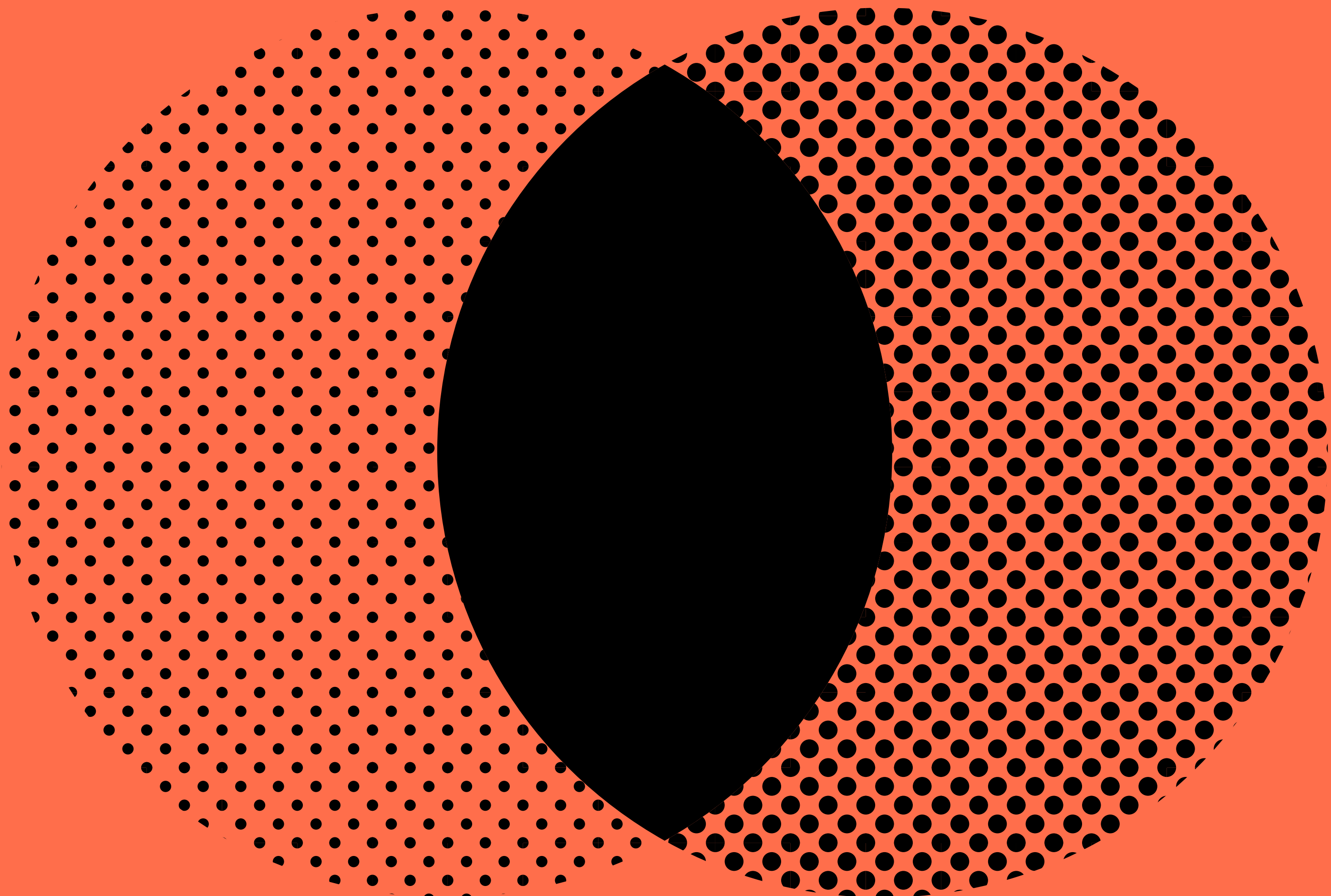
StudentSurvey.ie

PGR StudentSurvey.ie

Irish Survey of Student Engagement

Results of qualitative data analysis projects

Report 1 of 5



HEA | HIGHER EDUCATION AUTHORITY / AN tÚDARÁS um ARD-OIDEACHAS

iua | IRISH UNIVERSITIES ASSOCIATION

thea

USI — Union of Students in Ireland / Aontas na Mac Léinn in Éirinn

Results of qualitative data analysis projects

Name of Report: Report on the Analysis of Qualitative Data from StudentSurvey.ie (the Irish Survey of Student Engagement) 2016 - 2020

Authors: Dr. Stephen Erskine & David Harmon

Organisation/ Institution: Insight Statistical Consulting

Foreword

The StudentSurvey.ie Steering Group is pleased to publish the results of five research projects analysing the qualitative data generated by the free-text response questions in StudentSurvey.ie and PGR StudentSurvey.ie. The results contained within this report make up one part of this research series.

Five projects were funded by research bursaries offered by StudentSurvey.ie in October 2020. The aim of the bursary awards was to promote greater ownership and encourage wider use of the StudentSurvey.ie and PGR StudentSurvey.ie data. Proposals for the analysis of the qualitative data emerging from StudentSurvey.ie and PGR StudentSurvey.ie were invited from members of the research community within the participating institutions, as well as commercial data analysis companies. The projects were completed in May 2021.

Each project is an independent project undertaken by qualified and experienced researchers on behalf of StudentSurvey.ie. Each project took a unique approach. Some projects involved analysis of all the qualitative data for a given year(s), and some homed in on a subset of the data. Some undertook a qualitative methodology, while others applied quantitative methods to qualitative data. The commonalities between all five projects are that they all utilised well-grounded methodologies, offer mechanisms for replication of the analysis in future years, and are innovative and authentic.

These results are the first of their kind for StudentSurvey.ie and PGR StudentSurvey.ie and we hope they are the first of many research projects involving the qualitative results of these surveys.

What are StudentSurvey.ie and PGR StudentSurvey.ie?

StudentSurvey.ie (the Irish Survey of Student Engagement) is an annual national survey of student engagement among first year undergraduate, final year undergraduate and taught postgraduate students in higher education institutions in Ireland.

PGR StudentSurvey.ie (the Irish Survey of Student Engagement for Postgraduate Research Students) is a biennial national survey of student engagement among Masters by Research students and PhD students in higher education institutions in Ireland.

Both surveys are designed to focus on student engagement, namely the amount of time and effort that students put into meaningful and purposeful educational activities, and the extent to which institutions provide such opportunities and encourage students to engage with them. The data collected reflect students' self-reported perceptions of their experiences.

Report on the Analysis of Qualitative Data from StudentSurvey.ie (the Irish Survey of Student Engagement) 2016 - 2020

Authors: Dr. Stephen Erskine
David Harmon

Insight Statistical Consulting
60 Merrion Square
Dublin 2, Ireland
www.insightsc.ie

Date: April 2021

Contents

List of Tables

List of Figures

1. Introduction and Background

StudentSurvey.ie (the Irish Survey of Student Engagement) is an annual national survey of student engagement among first year undergraduate, final year undergraduate, and taught postgraduate students in 25 higher education institutions in Ireland. The survey is designed to focus on student engagement, namely the amount of time and effort that students put into meaningful and purposeful educational activities, and the extent to which institutions provide such opportunities and encourage students to engage with them. The data collected reflect students' self-reported perceptions of their experiences. Significant quantitative data are collected by survey from 67 multiple choice questions, and numerous student and course characteristic variables are also recorded.

In addition to the quantitative component, there are two qualitative questions within the survey which are open-text and ask:

1. What does your institution do best to engage students in learning?
2. What could your institution do to improve students' engagement in learning?

These questions have been in use since 2016, and the number of students who answer the survey has increased significantly over the years. In 2020, these questions were answered by approximately 26,000 respondents. Students may answer the survey in Irish or English, with the majority choosing to answer in English.

As such, since 2016 the two qualitative questions have generated an extremely large corpus of data which as a resource, for understanding students' thoughts and experiences in higher education, has been under-examined[1]. This under-examination of the material is perfectly understandable as the volume of material produced by these two questions each year is significant, and we would argue requires a different approach to standard qualitative analysis because of the sheer scale of the material.

[1] Irish Survey of Student Engagement. National Report 2020, p79.

For example, if one were to take the comments provided by students from 2016 to 2020 and place them into a standard A4 document, it would amount to almost 8,000 pages of A4 text. For an average reader this would take approximately 200 hours to read through once, and even longer if one wanted to code the material, identify themes, add comments, elaborate to set of generalisations, and collate all of the material found into a document to report the findings as a sequence of tasks necessary to conduct qualitative research[2]. Thus, a key point to be noted at the outset of this report which will be returned to numerous times throughout is the amount of material at hand for analysis here.

With this amount of data, steps have to be taken to handle the corpus of information in a suitable manner as hand-coding of material becomes too unwieldly to be practical. As such, this report has combined researcher-led coding of the material with computer-assisted content analysis methods to draw out key information on what students think about student engagement in their higher education institutions. In addition, this report provides a framework for further research to build upon as our analysis utilised the open-source statistical software *R* and its graphical interface R Studio, along with a suite of R packages. Because of this, our project is replicable, and our coding framework and code can be extended to further iterations of the Irish Survey of Student Engagement.

1.1 Methodological Approach

As the data examined in this study are qualitative in nature, a case could be made that it requires qualitative analytical methods to provide us with answers. However, we would argue that the debate around qualitative versus quantitative methodologies is a false dichotomy and the methodology used in research should follow naturally from the requirements of the research. As Ken Benoit argued[3]:

> The move from qualitative to quantitative measurement occurs as more information is incorporated. It is also a natural consequence of any effort to

[2] Robson, C. Real World Research. (Blackwell Publishing, 2002), p459.
[3] Benoit, K. How Qualitative Research Really Counts, in Qualitative Methods Newsletter (Spring), 2005.

> compare observations. Comparison implies ordering, whether on a qualitative or quantitative dimension. Ordering implies by nature that one quality is stronger or greater in one observation than in another. And relations such as "stronger" or "greater" imply, whether this is made explicit or not, a relative degree of quantity, even if the characteristic being compared is discussed in purely qualitative terms. The act of comparison, therefore, naturally and readily lends itself to quantification.

In this research project, we have two questions for which each has around 90 thousand responses. It is possible to examine individual students and their responses, but this naturally gravitates to how many students have responded, to ultimately how students **in general** have responded. The use of 'in general' marks a subtle shift from qualitative to a quantitative methodology already.

Though of course, this is not to say that qualitative analysis formed no part of the research. Rather, to begin the project we needed familiarity with the data and we had to immerse ourselves in the material thoroughly in order to get a grounding in the sentiments contained in the open-ended responses. We also read each of the individual reports from each round of the survey to provide context and information on what was found in the quantitative analysis[4].

Once we had familiarity with the data, we could code the data using statistical software and use this to separate and draw out the core and latent themes present in the corpus. Qualitative analysis can get us only so far by itself, much like using only quantitative analysis without familiarity with the material would have left us operating in a vacuum without key contextual information. As such, supplementing our initial grounding in the material with quantitative text analysis, in our view, only increases the validity of the overall findings and provides a greater depth of understanding the material at hand.

[4] The Irish Survey of Student Engagement (ISSE). Results from 2016.
The Irish Survey of Student Engagement (ISSE). Results from 2017.
The Irish Survey of Student Engagement (ISSE). Results from 2018.
Irish Survey of Student Engagement. National Report 2019.
Irish Survey of Student Engagement. National Report 2020.

The rest of this report details the analysis conducted and begins in the next section with an overview of the demographics of students that provided responses to the two open-text questions. Chapter 2 then examines student responses to the question: “What does your institution do best to engage students in learning?”. This is followed in Chapter 3 by an investigation into student responses to the question “What could your institution do to improve students' engagement in learning?”. Chapter 4 provides some of the conclusions that can be drawn from the analysis.

1.2 Meta-analysis of the Survey Responses

Before beginning the analysis of the individual questions, it is worth examining who is providing material in the open text fields and seeing if these respondents differ from the overall population of respondents. Ideally, we would like for any inferences reached in the next two chapters to be equally valid for all students, and for this to be the case we would want the population that provided comments to not be too different from the overall student population.

Table 1.1: Overview of survey responses by year

Year of survey fieldwork	Provided a comment?		Total
	No	Yes	
2016	12,894	16,279	29,173
2017	16,069	19,781	35,850
2018	17,892	20,479	38,371
2019	18,746	21,812	40,558
2020	17,394	27,313	44,707
Grand Total	**82,995**	**105,664**	**188,659**

Table 1.1 presents the number of survey responses each year, and the number of students that provided a comment in at least one of the two open-text fields. As can be seen from this

table, the number of students providing a response increases each year, and the number of students providing a comment also increases over time.

Table 1.2 provides this breakdown of students providing responses and comments to the open-text fields across time in percent, and for four of the five years the percentage providing a comment was around 55 percent. In 2020, this increased to 61 percent. Table 1.1 showed that the number of responses and the number of respondents providing comments in the open-text fields were increasing over time, though as can be seen in Table 1.2, as a share of the total the proportion remains broadly similar year-on-year.

Table 1.2: Percent of survey responses providing response to open-text questions

Year of survey fieldwork	Provided a comment?	
	No	Yes
2016	44%	56%
2017	45%	55%
2018	47%	53%
2019	46%	54%
2020	39%	61%
Grand Total	**44%**	**56%**

Within the students who provided a comment, one question that arises is; are respondents who provide a comment significantly different from those who do not provide a comment? Evidence from other studies tend to show that older, and female students are more likely to fill in a survey than other groups. Typically results are then weighted to accommodate these differences and ensure that the sample is representative of the whole population. Unfortunately, this method is unavailable to us for qualitative data because we cannot weight open-text responses to fit the known student population in the manner we would weight cases to ensure the demographics of respondents fit the overall student population.

As such it is necessary to identify at this preliminary stage if there are any discrepancies between the sample and the overall population. Table 1.3 shows the proportions of respondents providing a comment by gender. It would not be desirable to see a large discrepancy in the percentage of female (or male) students providing a comment. This table shows that we are not seeing an over representation of female respondents as the percentage providing a comment matches the male proportion quite well.

Table 1.3: Percent of survey responses providing response to open-text questions by gender

Year	Gender	Provided a comment?		Total N
		No	Yes	
2016	Female	44%	56%	17,208
	Male	44%	56%	11,965
2017	Female	46%	54%	20,845
	Male	44%	56%	15,005
2018	Female	47%	53%	22,743
	Male	47%	53%	15,628
2019	Female	46%	54%	23,841
	Male	46%	54%	16,709
2020	Female	39%	61%	26,342
	Male	39%	61%	18,330
Grand Total		**44%**	**56%**	**188,659**

Table 1.4 shows the proportions of respondents providing a comment across two age bands, and again we would not want to see a large discrepancy in one group of students providing a comment over the other. In this table, there does appear to be some evidence of an age effect year-on-year, as older students appear to be more likely to have provided a comment than younger students, but again this is not considered too large to be of concern to the overall results.

Table 1.4: Percent of survey responses providing response to open-text questions by age

Year	Age	Provided a comment?		Total N
		No	Yes	
2016	23 years and under	46%	54%	18,448
	24 years and over	40%	60%	10,682
2017	23 years and under	46%	54%	23,454
	24 years and over	42%	58%	12,396
2018	23 years and under	48%	52%	25,027
	24 years and over	44%	56%	13,344
2019	23 years and under	47%	53%	26,708
	24 years and over	44%	56%	13,850
2020	23 years and under	40%	60%	29,717
	24 years and over	37%	63%	14,990
Grand Total		**44%**	**56%**	**188,659**

Finally, for the purposes of the analysis, responses provided in Irish have been translated into English. Doing this allows for unified analysis of all the data, rather than separate analyses. We have, however, included a variable in the analysis which captures the language of the original responses, as an essential part of the analysis would be to see whether students who use Irish in their day-to-day life in higher education have different experiences from students who primarily use English. The table shows that a similar percentage choose to complete the survey in Irish (and English) each year.

Table 1.5:Distribution of survey responses in Irish and English (Count and Percent)

	Year				
Language	**2016**	**2017**	**2018**	**2019**	**2020**
English	16,131 (99%)	19,619 (99%)	20,294 (99%)	21,689 (99%)	27,100 (99%)
Irish	148 (1%)	162 (1%)	185 (1%)	123 (1%)	213 (1%)
Total N	**16,279**	**19,781**	**20,479**	**21,812**	**27,313**

Overall, there appears to be no evidence at this initial stage to suggest that respondents providing comments deviate substantively from the overall student population. As such, we can have confidence that the results drawn from the next steps in the analysis are applicable to the whole student population. The first stage of which is covered in the next chapter.

2. “What does your institution do best to engage students in learning?”

This chapter of the report covers the analysis undertaken to provide some answers to the open-text question posed to students; “What does your institution do best to engage students in learning?” Which for the shorthand purposes will hereby be referred to as Q1.

As noted already, the sheer volume of data provided by students in their responses to this question is the key difficulty in evaluating the content of students’ responses. For example, in the datafile provided by StudentSurvey.ie, the content of Q1 from 2016 to 2020 when inserted into a standard A4 document, runs to over 4,000 pages. As such, with this amount of information it becomes all too easy to be overwhelmed unless a strategy to break the data down into more manageable units of analysis is utilised.

This chapter covers the analysis conducted to break this material into more readily comprehensible units first, and then discusses the subsequent analysis which reintegrates these units to gain a systematic understanding of students’ responses which can then be used to illuminate areas in which HEIs are succeeding in engaging students.

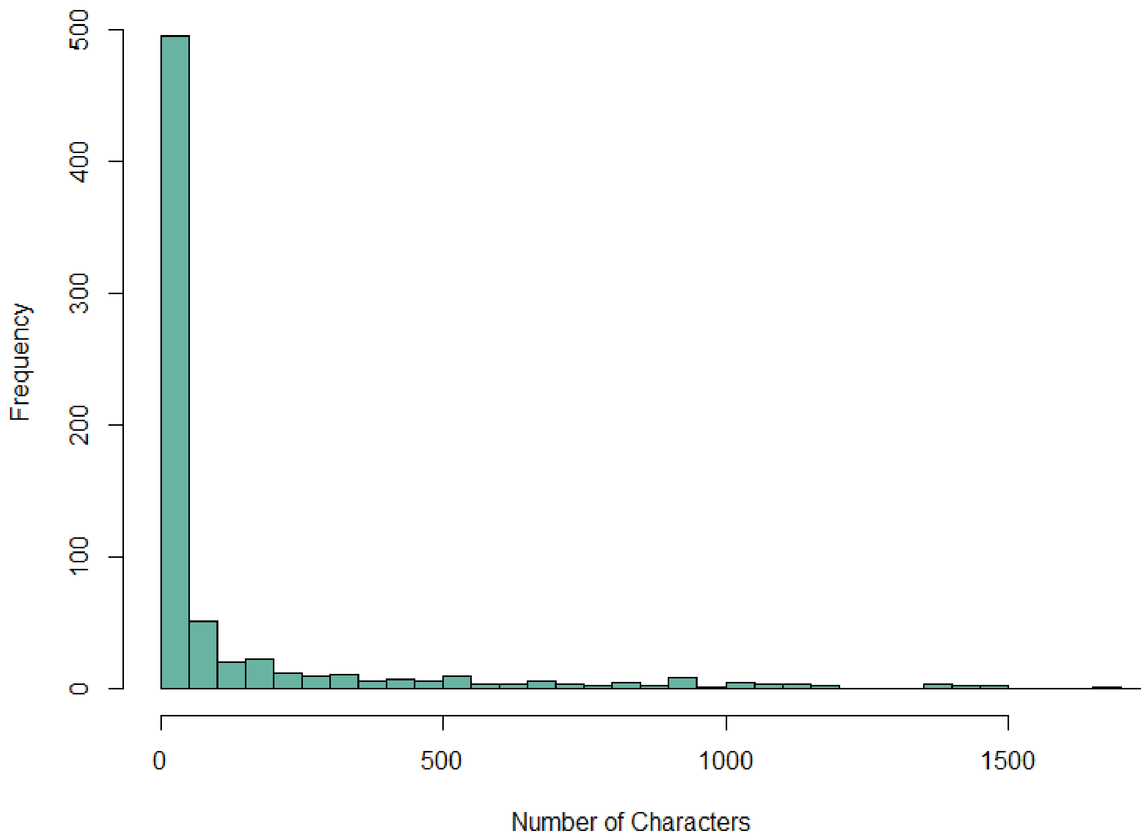
2.1 Meta-analysis and Frequency of Characters

To begin, the question asked in this portion of the survey was open-text and had no limit on the amount of material that respondents could provide. A result of this is that there is considerable variation in the amount of text students provided. Some provided terse answers, other miniature essays. Table 2.1 provides some summary statistics of the number of characters used by students in their responses, that is the number of letters, spaces and punctuation marks provided in a response. This table shows that on average, the length of a response was close to 70 characters in length, though the median length was only 47 characters. The minimum length of a response was three characters, which is discussed further below, and the maximum length was 1,668 characters.

Table 2.1: Summary statistics of the number of characters in Q1

Mean	Median	Standard deviation	Interquartile Range	Minimum	Maximum	Total N
69	47	75	67	3	1,668	97,804

Figure 2.1: Distribution of the number of characters in students' responses to Q1

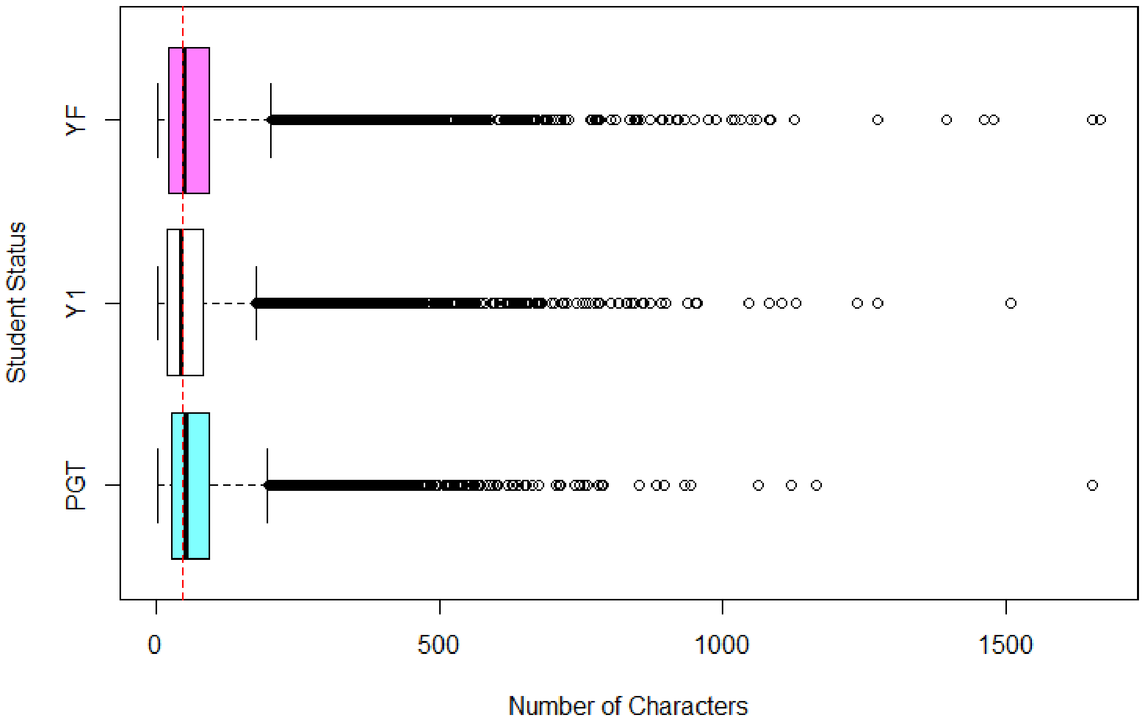


A large difference in mean and median values is typically an indication of a non-normal or skewed distribution, and Figure 2.1 supports this. Over fifty percent of responses use fewer than 50 characters, whereas there are a small number of cases which use a lot of characters. A distribution such as this pulls the mean away from the median.

While this distribution is interesting, it is not wholly unexpected. It would only really be of a concern to the analysis if certain groups of students had different patterns of responding to the question. For example, if first year students provided one word answers whereas in contrast taught postgraduate students provided comprehensive itemised lists. To ensure that this is not the case, Figure 2.2 provides three boxplots of the distribution of the number of characters in students' responses across the three categories of students surveyed (first year (Y1) and final year undergraduates (YF) and taught postgraduates (PGT)).

As can be seen from this chart, this distribution of each follows the skewed distribution shown in Figure 2.1. However, each group also has a similar median shown in the black central bars in the boxes of the boxplots and each appear to be close to the overall median, which is presented as a red dashed line in the graph. Each also has a similar range, so from this perspective there does not appear to be any indication that different groups of students tend to provide longer or shorter responses than other groups of students.

Figure 2.2: Boxplots of the distribution of characters in responses to Q1 by Student Status



From a meta-perspective, the number of characters that students use when answering tells us something about the variation in the amount of information students are willing to provide, but it tells us nothing about the information itself. To disaggregate this and begin to understand the information provided we need to be reasonably confident in the quality of the material to begin with.

This confidence in the quality of the material was achieved by taking the data through a series of steps which changed it from a 'raw' format to being 'cleaned' to ensure that all the data contained is treated equally by the statistical software. A copy of the original Q1 variable is kept throughout as a reference and quality check.

The first step in this case was gaining a familiarity with the data by reading over many of the comments. This was very time-intensive but invaluable in understanding how students approached answering the question at hand. As shown in Table 1.5 above, a minority of students provided their answers in Irish but for these answers to be analysed in the same way as all other comments these had to be translated into English. We did, however, included an additional variable in the analysis which captured the language of the original response, as an essential part of the analysis was to see whether certain student groups like those who use Irish in their day-to-day life in higher education have different experiences from students who primarily use English.

This step was conducted in tandem with a comprehensive spell check which tried to ensure that any common incorrect spellings present in the data were rectified and that any common differences were changed to a single keyword. For example, a common answer to Q1 was some variant of "the student union" however there were a myriad of ways that students provided this response such as "the su", "the SU", "the <name of their HEI> SU", "the students union", "the students' union", "the student's union" and other combinations thereof. Because of this, all these variants were changed to "student_union" so that they would be identified as being all of one group by the statistical software. The underscore symbol was used to identify that this is a keyword which tends to be separated by a space, but which should be evaluated together in the analysis as being different from the individual words "student" and "union". This is discussed further below as a number of items of interest in the analysis are multiword phrases which have been compounded to ensure that they are correctly identified as being distinct from their constitutive components and are not analysed separately.

Tables 1.1 and 1.2 show that the majority of students in each year of the survey provided an answer to either of the open-text questions; the vast majority of these comments were coherent answers which were left untouched by the cleaning process. A reasonable proportion were corrected for spelling errors as noted above. However, a small proportion of the text for Q1 was unusable for a variety of reasons. Firstly, the longer the text provided by a respondent the easier it is to gain some comprehension about the intent of the student even if their original text was corrupted in some way (with a spelling error for example). However, under a certain length it was impossible to ascertain what the respondent was presumably trying to say. In this analysis, comments of one or two characters fell into this category. Secondly, a similar problem was evident with comments like "Yes" or "No", as the intent of these remarks could not be accurately determined. Thirdly, there were a number of random placeholder text strings which had been entered by students which did not actually contain any content, for example, "asdf" or "gygygygy". Finally, some text contained expletives, and raises the question of whether these should be taken at face value or incidences of sarcasm or irony. As disentangling this and the other unusable text could not be adequately resolved, all of these cases were removed from the analysis by being coded as unusable and filtered out from the subsequent analysis.

Once the preliminary stages of cleaning the corpus of data had been completed, the data were ready to be entered into the statistical software[5]. The first step of this was to remove punctuation and symbols that cluttered the corpus. The second step was to segment the corpus of data into individual tokens, usually words separated by white space. However, as noted above, through immersion in the material from the outset, we noted some words which tended to be associated with one another. To avoid these being lost to the analysis, underscores were inserted between these words, so that these multiword phrases would be compounded and seen by the software as being one unit. These phrases were: "continuous_assessment", "academic_staff', "small_classes", "real_world" "real_life", "group_projects", "practical_work", "guest_speakers", "case_studies", "teaching_assistants", and "student_union".

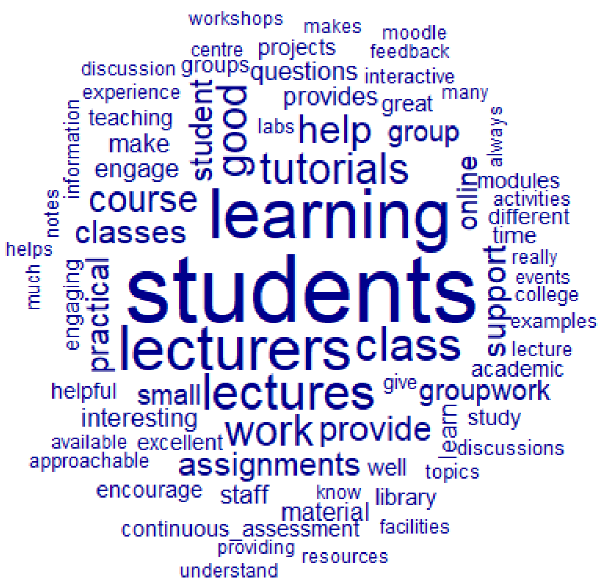
[5] Throughout the analysis, we used a combination of R, its graphical user interface RStudio and a set of text analysis packages to conduct the analysis. All of this material is provided separately in a set of R Scripts.

The final step was to remove common 'stop words' that are of little intrinsic value to the analysis (such as "the", "a", "an", and "in") and whose presence would suppress words and phrases that are important to the analysis. The *Quanteda* R package has a built-in list of these common words, though this list was also supplemented further by examining the frequency of individual words and removing extraneous and unnecessary words by adding them to the stop word list. A full list of the stop words removed is provided in Appendix A.

Traditional text analysis would also stem words removing suffixes and common endings of words so that word stems could be examined together. This has not been done here as to do so would lose nuance and mean that we would be unable to distinguish between words that are very similar but have crucial differences. For example, lectures and lecturers; tutorials and tutors; practical and practicals.

At this point we are left with a corpus of data that is now ready for analysis. One initial way of examining the corpus would be to view the frequency of keywords. Figure 2.3 presents the 75 most frequently occurring words in a word cloud, with the relative size of the word being an indication of its frequency. As we can see from this graphic, "students", "learning", and "lecturers" are the most commonly used words in the corpus.

Figure 2.3: Word cloud of the 75 most frequently used words

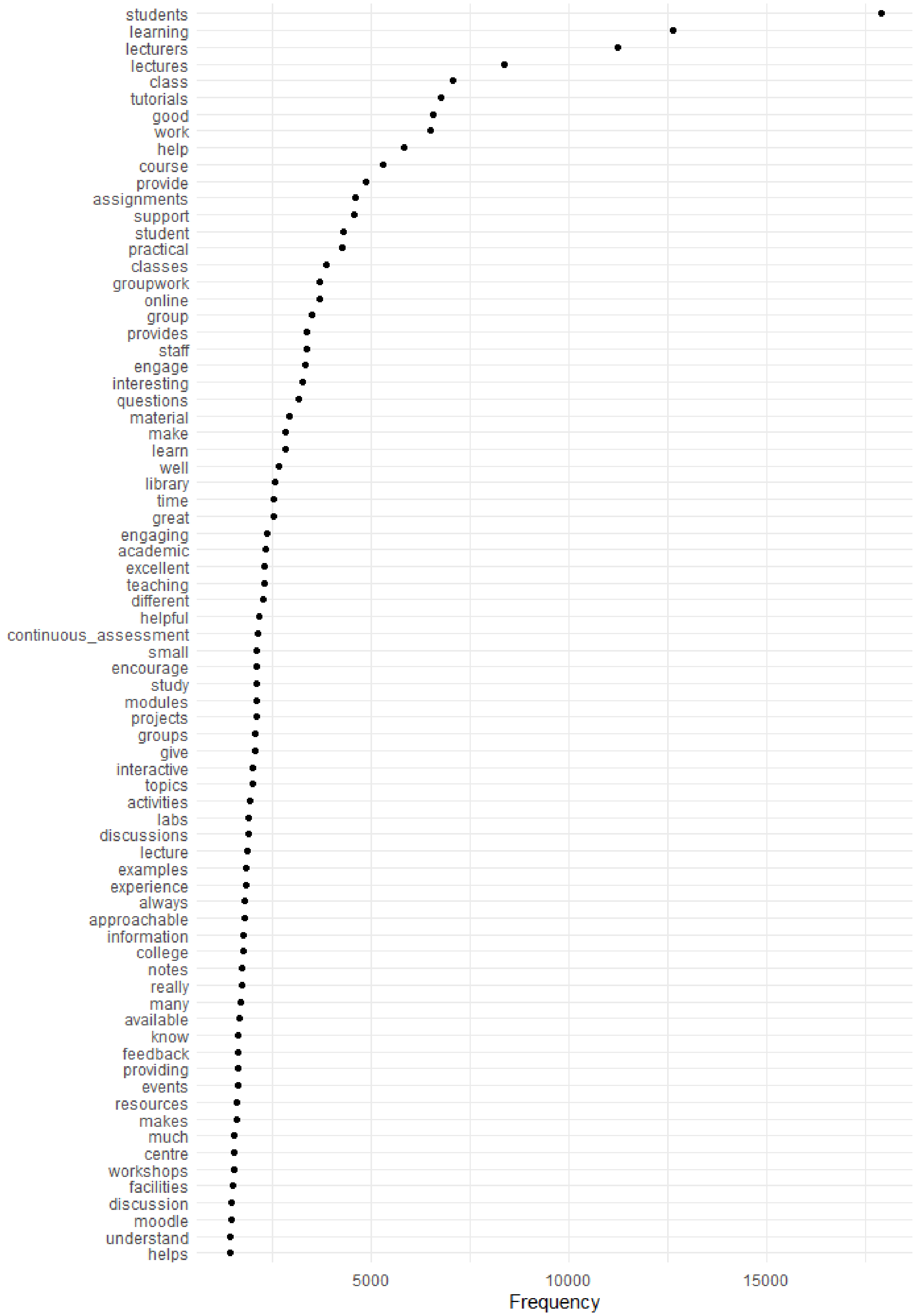


However, by themselves word clouds are not particularly enlightening. They are a good way of viewing the corpus quickly but beyond the most frequently used words it is difficult to evaluate the relative placement of terms. For example, “workshops”, “makes”, and “moodle” are at the top of the graphic, but which one is the most used of these terms? As such, this word cloud should be seen as a useful first step but not an end point in itself.

Figure 2.4 uses the same information as Figure 2.3 but plots the relative frequency of the top 75 words instead. From this we can see that “students” is the most used word with over 17,500 instances in the corpus. After this there is a large gap between the most used and second most used with “learning” being used around 12,500 times. In third place, close to “learning” is “lecturers” with around 11,000 instances. The chart also shows that the most frequently used words are used a lot. For example, all of the top 10 words have over 5,000 instances in the corpus. In contrast, the bottom 45 words in the chart have less the 2,500 instances each.

There are 14,300 unique words used in the clean Q1 corpus, and these words are used 592,000 times in total. However, the top 75 words account for over forty percent of the words used in the Q1 corpus.

Figure 2.4: Relative frequency of the top 75 most frequently used words (Q1)


students
learning
lecturers
lectures
class
tutorials
good
work
help
course
provide
assignments
support
student
practical
classes
groupwork
online
group
provides
staff
engage
interesting
questions
material
make
learn
well
library
time
great
engaging
academic
excellent
teaching
different
helpful
continuous_assessment
small
encourage
study
modules
projects
groups
give
interactive
topics
activities
labs
discussions
lecture
examples
experience
always
approachable
information
college
notes
really
many
available
know
feedback
providing
events
resources
makes
much
centre
workshops
facilities
discussion
moodle
understand
helps
5000
10000
15000
Frequency

While we could view the corpus as being a single body of data collected over five years (2016-2020), it can also be viewed as five separate collections of data, each collected in a separate year, and these corpora are never going to be identical, as each year brings with it a different cohort of respondents with different experiences, along with some changes in approaches to teaching and learning by their HEIs even if the curriculum remains the same[6].

However, it is worth examining the popularity of words in the corpus over time to see if there are any substantial changes or if the popularity of words used by respondents remains stable over time. Table 2.2 presents the ranking of the top 100 words used in each year and has been colour-coded so that high ranked words are green, and lower ranked words are red, with words in-between these poles moving along a continuum from yellow to orange. Changes in the colours over time indicate shifts in ranking, so moving from green towards red means that the word has fallen down the ranking. Conversely, moving from red to green indicates increased usage and a rise in its relative rank.

As can be seen from this table, at the top end of the table there is remarkable stability. Beyond this, there appears to be some natural variation across years with most words moving up or down somewhat arbitrarily. Notably though, there are a few words that rise through the rankings across the years. For example, "activities" moves from 63rd position in 2016 to 32nd position in 2020, "continuous assessment" moves from 65th position in 2016 to 31st position in 2020, and "interactive" moves from 83rd position in 2016 to 29th position in 2020. These shifts appear to indicate that because these are being more often mentioned by respondents, these facets of respondents' experiences are becoming more important over time.

It is worth pointing out though that these may be becoming more important, from this chart we remain unable to say if this is a (broadly) positive or negative for students. Students may be mentioning each of these more in 2020 than in 2016 but they may all be saying that they hate continuous assessments, dislike interactivity, and do not see the purpose of activities. Thus, we should be careful about interpretation here. This point is addressed further below.

[6] We are also implicitly assuming here that each year has been provided with broadly similar experiences, which may not be the case in the 2021 iteration of the survey. For example, between 2016 and 2019 there were no mentions of 'covid' or 'coronavirus'. In 2020, 'covid' was mentioned twice and 'coronavirus' once. Due to the events post-data collection in 2020 and in 2021, we can expect these keywords to form a considerable component of the corpus in 2021.

Table 2.2: Ranking of the top 100 most frequently used words by year

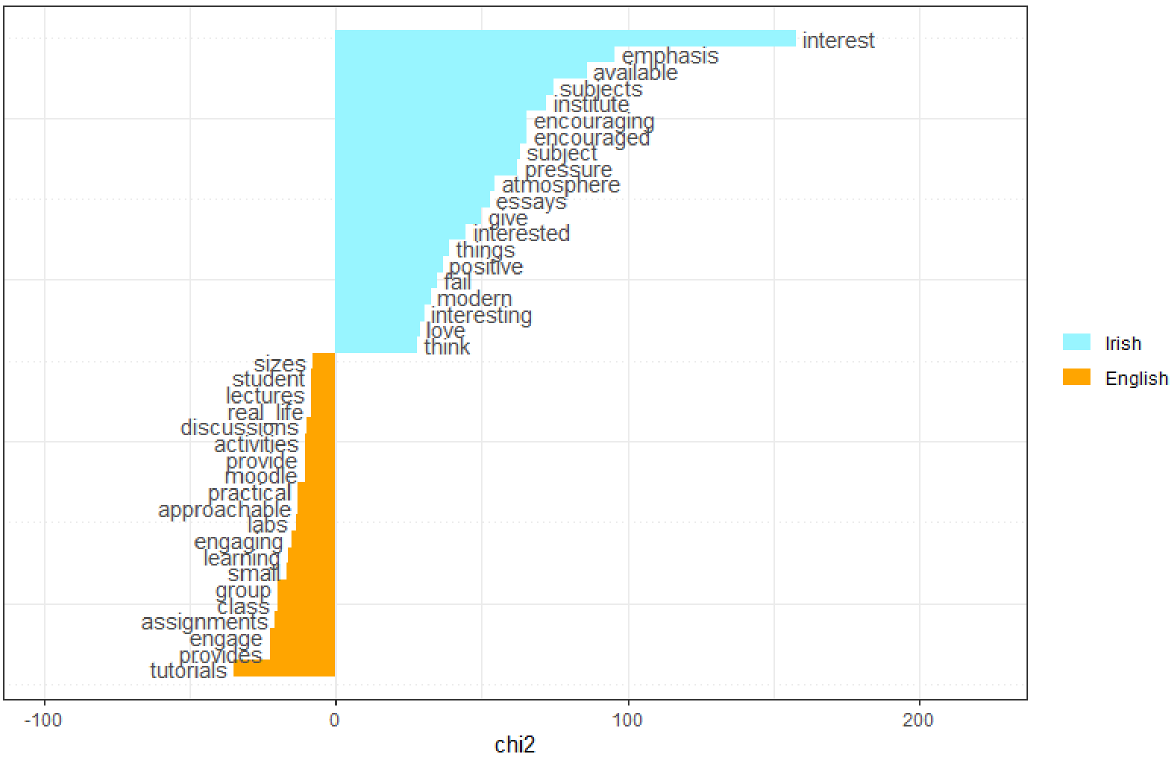
Keyword	2016	2017	2018	2019	2020	Keyword	2016	2017	2018	2019	2020
students	1	1	1	1	1	topics	51	48	36	49	50
learning	2	2	2	2	2	information	52	53	56	60	52
lecturers	3	3	3	3	3	many	53	52	68	58	64
lectures	4	4	4	4	4	discussions	54	62	46	46	47
class	5	5	5	5	5	labs	55	67	50	45	39
work	6	6	7	7	8	modules	56	42	40	42	42
tutorials	7	8	6	6	6	projects	57	57	44	36	33
good	8	7	8	8	7	encourages	58	63	75	83	97
help	9	9	9	9	11	know	59	61	62	63	66
course	10	10	10	11	10	makes	60	51	64	73	71
student	11	13	15	17	19	examples	61	58	49	51	44
provide	12	11	11	13	12	providing	62	77	63	57	57
support	13	12	12	14	16	activities	63	59	53	48	32
assignments	14	15	14	12	9	understand	64	72	73	85	75
practical	15	17	16	15	15	continuous_assessment	65	49	37	44	31
staff	16	18	24	24	24	much	66	64	76	68	65
provides	17	16	23	23	27	workshops	67	88	84	55	63
classes	18	14	13	10	13	year	68	84	90	98	116
online	19	26	21	20	14	best	69	60	100	74	73
interesting	20	20	22	25	22	friendly	70	83	96	96	84
engage	21	19	19	22	23	environment	71	80	77	90	82
small	22	21	18	18	20	resources	72	73	52	69	59
learn	23	23	28	27	36	feedback	73	75	60	52	60
material	24	30	26	26	25	think	74	82	105	127	96
well	25	28	31	31	34	discussion	75	86	67	77	58
group	26	25	20	19	18	institution	76	81	89	89	101
time	27	31	29	38	40	facilities	77	65	59	72	77
excellent	28	34	41	40	43	extra	78	79	86	70	83
groupwork	29	22	17	16	17	moodle	79	66	88	75	62
great	30	29	34	33	35	interaction	80	107	92	93	94
make	31	24	27	29	26	helps	81	69	70	71	79
questions	32	27	25	21	21	events	82	55	72	64	55
library	33	33	30	28	30	interactive	83	70	48	32	29
helpful	34	36	47	39	54	feel	84	74	93	105	99
teaching	35	37	38	41	38	exams	85	92	109	102	90
academic	36	38	33	35	37	tutors	86	85	81	76	85
college	37	44	71	66	70	blackboard	87	76	66	95	114
approachable	38	43	55	62	76	people	88	71	79	78	78
study	39	35	43	47	49	easy	89	89	114	107	105
experience	40	46	65	56	61	open	90	108	82	80	80
different	41	32	39	34	41	centre	91	78	54	59	67
encourage	42	40	32	50	48	subjects	92	95	108	136	139
always	43	47	57	54	68	writing	93	97	74	79	103
give	44	41	45	43	45	outside	94	102	87	125	112
available	45	68	51	65	69	presentations	95	96	83	82	81
engaging	46	39	35	30	28	understanding	96	132	113	109	135
lecture	47	50	58	53	51	better	97	94	116	126	126
notes	48	56	61	67	56	need	98	99	123	110	143
groups	49	45	42	37	46	academic_staff	99	116	147	129	127
really	50	54	69	61	53	lecturer	100	101	80	99	106

2.2 Relative Frequency Analysis (Keyness)

We can also compare the whole corpus across student status, and the language respondents used to complete the survey, to see if these various groups use different terminology to each other. To identify significant differences across groups contained within a corpus we use a statistical measure called *keyness* which uses the relative frequency of words across two parts of a corpus to see if there are differential associations of keywords between a target and a reference group.

For example, Figure 2.5 presents the relative frequencies of words used by students completing the survey in Irish against those completing the survey in English.

Figure 2.5: Relative frequency analysis (keyness) by language



Students completing the survey in Irish use the word ‘interest’ significantly more than students conducting the survey in English. On the other side, students completing the survey in Irish tend to use the terms “tutorials”, “provides”, “engage” less than students conducting the survey in English. Though as shown in Table 1.5, the Irish corpus is very small compared against the English corpus, so we are likely to see some differences purely due to the fact that because the Irish corpus is smaller, it is less likely to contain the range of words present in the English corpus.

It is also feasible that the status of a student may have an influence on their overall outlook, for example, a first year undergraduate is less likely to be worrying about their dissertation due in a few years' time than a final year student who has begun drafting their dissertation, and this should have a concurrent impact on the words and phrases they use in the survey with final year students being more likely to mention dissertations or their thesis than first years. In short, priorities for students may look very different if they are in their first year as compared with their final year, and postgraduate students are likely to have different experiences to those of undergraduate students.

Figure 2.6: Relative frequency analysis (keyness) by student status

First year undergraduates	Final year undergraduates	Taught Postgraduates

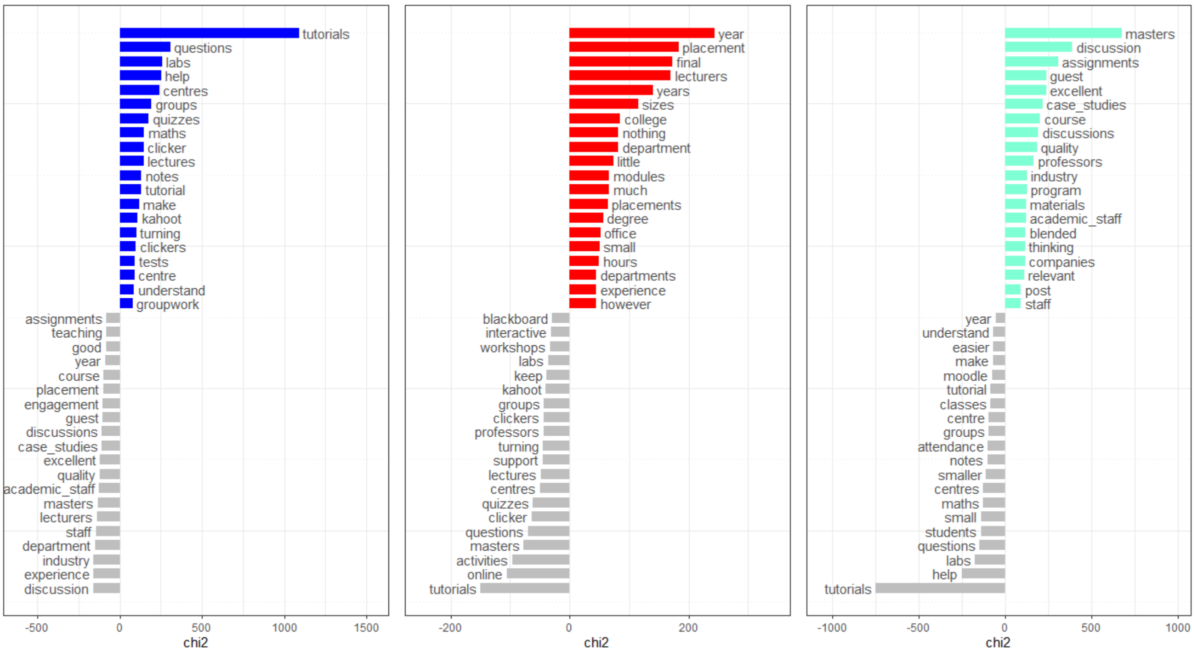


Figure 2.6 examines this by comparing each group against a reference group (i.e. the other two groups) to see the relative frequencies of keywords. The bars in grey are terms frequently used by the reference groups and those in colour are the target group for that chart.

From the first chart on the left we can see that first year undergraduates mention tutorials more than final year or taught postgraduates, potentially because these are often very new to them and are different from their prior experiences in second-level education. For final year undergraduates we can see that “final” and “year” are mentioned a lot, along with “placement” potentially referring to work experiences or placements which are presumably of higher priority to students facing the completion of their degree and many of which are likely to join the labour market post-completion of their degree. Taught postgraduates understandably mention “masters” more than undergraduates.

From a statistical perspective, chi-squared tests are sensitive to the size of the corpora we are testing, and we have a very large corpora with over ninety thousand individual comments, which can further be broken down into individual words. In essence, when we have large corpora, we are likely to generate a lot of keywords with high keyness values. As Gries observes, “many contemporary corpora provide basically guarantee that even minuscule effects will be highly significant”[7].

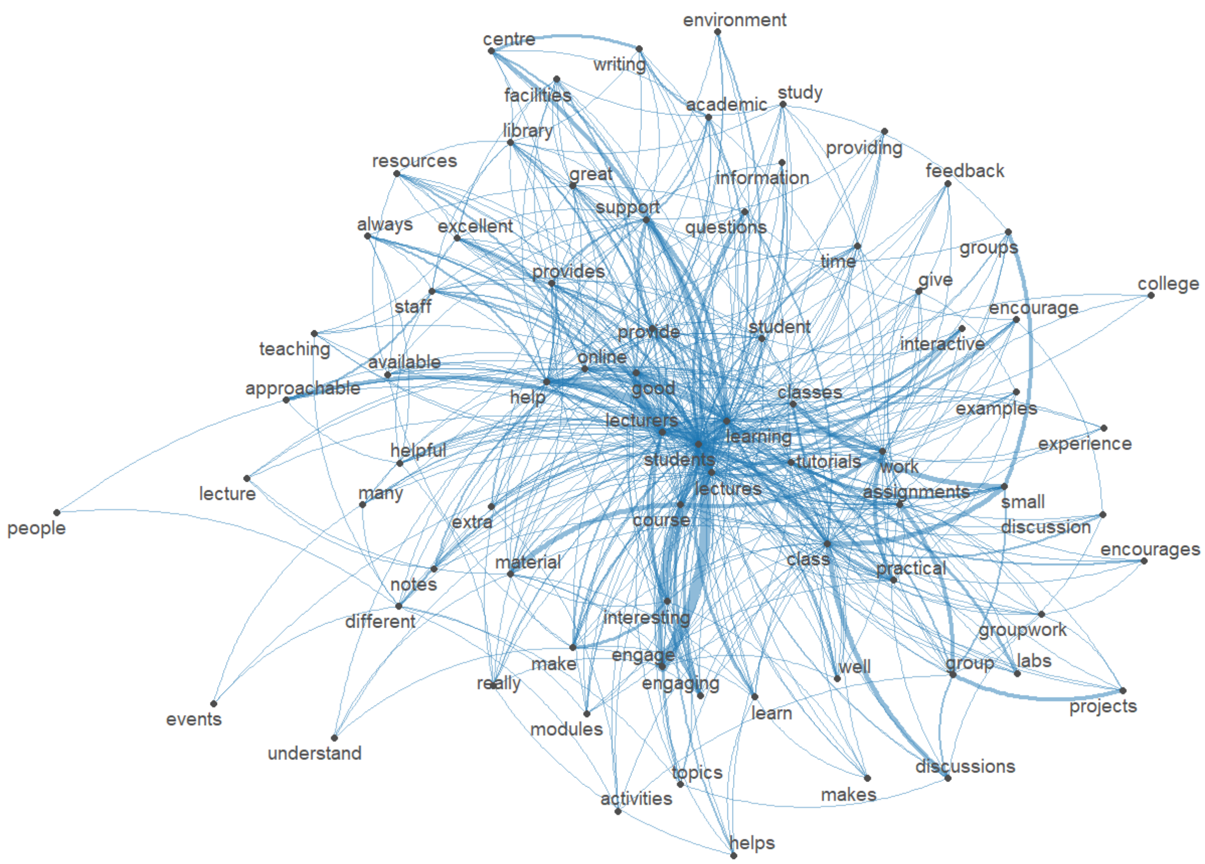
This brings up the question of whether these statistically significant effects are *substantively* significant. All of the words presented in Figures 2.5 and 2.6 are significantly different from the reference groups which lends support to the notion that these groups have subtly different priorities which are expressed through the different frequencies of certain words used, however, these charts do not show huge substantive discrepancies across groups; each groups experiences of student life have more in common with each other than their relative differences.

So far, the analysis has been of individual words, which has shown us the frequency of words used by students at an aggregate level, and some patterns with how these frequencies change when student groups are disaggregated or across time. The next step is then to identify which words are most associated with one another. Within the statistical software this was done by creating a feature co-occurrence matrix which records the number of co-occurrences of tokens.

[7] Gries, S. (2010). Useful statistics for corpus linguistics. Available from: https://www.researchgate.net/profile/Stefan-Gries-2/publication/267242111_Useful_statistics_for_corpus_linguistics/links/5466341a0cf2f5eb180168f4/Useful-statistics-for-corpus-linguistics.pdf

This feature co-occurrence matrix can then be visualised in a semantic network to illustrate which words are most associated with one another. The width of the bars linking words indicates the strength of the relationship between the words. Figure 2.7 presents a semantic network and as we would expect the three most frequently used words in the corpus thus far, “students, “learning”, and “lecturers” form the central axis from which all the other words branch.

Figure 2.7: Semantic network of feature co-occurrence matrix



2.3 Latent Semantic Scaling

While a semantic network is an appealing way of quickly visualising the corpus, it can be difficult to disentangle, especially if you have a specific interest in the words associated with one facet. As such, an alternative way of visualising the association between words is to use latent semantic scaling. Latent Semantic Scaling (LSS) is a flexible and cost-efficient semi-supervised document scaling technique.

This method has a couple of distinct advantages which the charts below demonstrate further. First of all, the scaling allows us to plot the frequency of words associated with each keyword of interest. Secondly, we can combine this with sentiment analysis through applying a sentiment dictionary to the corpus. This sentiment dictionary identifies words with specific positive and negative connotations which then allows us to identify the most used positive and negative words associated with key words of interest[8].

Figure 2.8 presents the words most commonly associated with 'lectures' and highlights the words with positive and negative connotations. As can be seen from the charts, there are a lot of positively associated words for lectures including interesting, good, helpful, and so on. There does not appear to be any negative words of note in the second chart.

[8] While this section discusses a few keywords of interest, the R code provided separately has the ability to apply the latent semantic scaling to any keyword of interest.

Figure 2.8: Positive and negative words associated with 'Lectures'

"Lectures"	
Positive	**Negative**

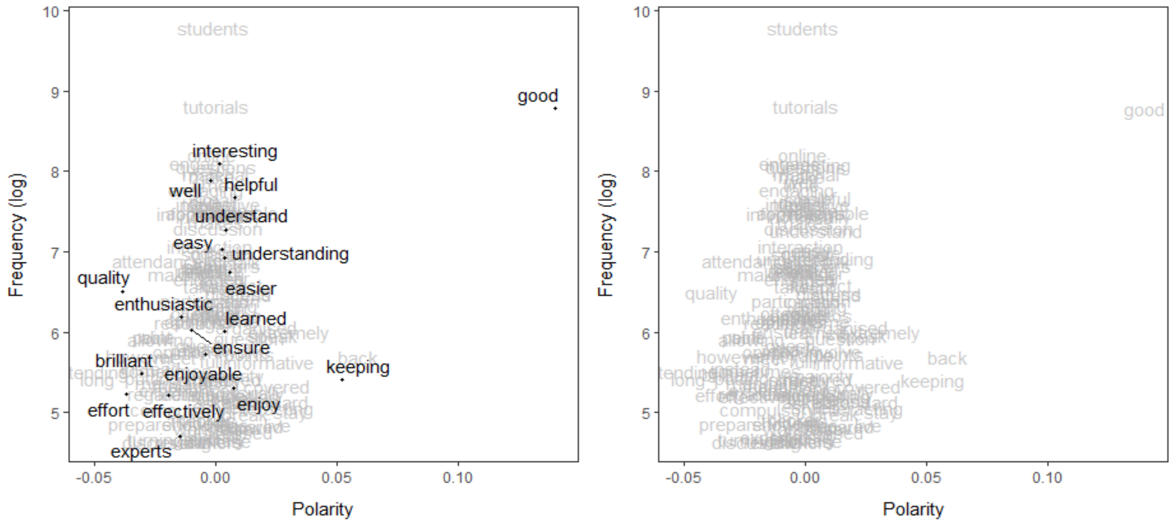


Figure 2.9 presents the words most commonly associated with 'lecturer' and highlights the words with positive and negative connotations. As can be seen from the charts, there are again a wealth of positively associated words for lecturer. The negative words of note, appear to be relatively infrequent in the corpus given their placement on the y-axis.

Figure 2.9: Positive and negative words associated with 'Lecturer'

"Lecturer"	
Positive	**Negative**

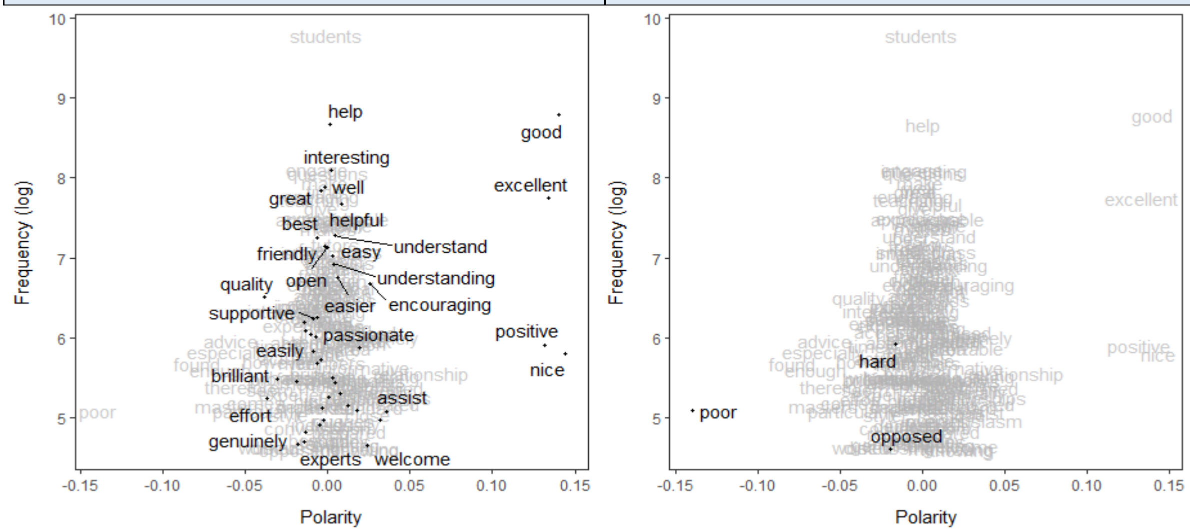
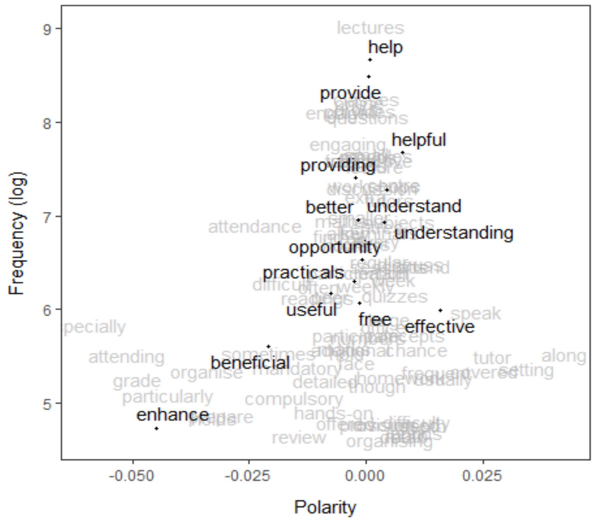


Figure 2.10 presents the words most commonly associated with ‘tutorials’ and highlights the words with positive and negative connotations. As can be seen from the charts, there are a wealth of positively associated words for tutorials with those that note that tutorials help or are helpful being used relatively frequently. The negative words of note appear to emphasise the difficulty some students have with the tutorial format.

Figure 2.10: Positive and negative words associated with ‘Tutorials’

“Tutorials”	
Positive	**Negative**



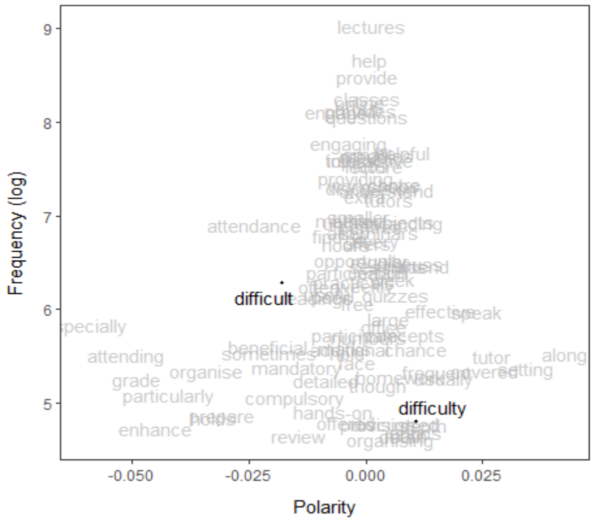
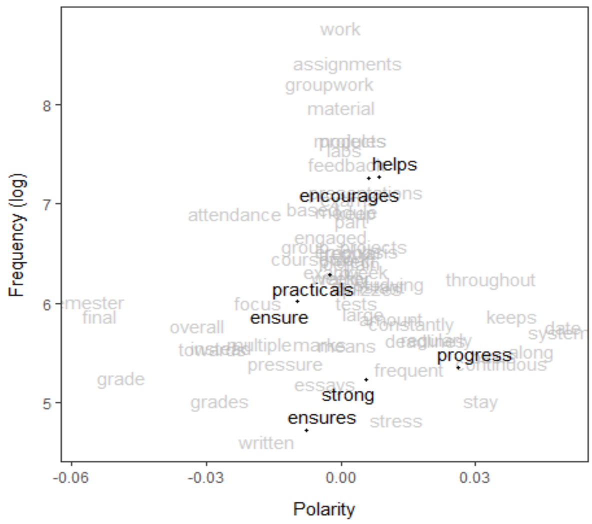
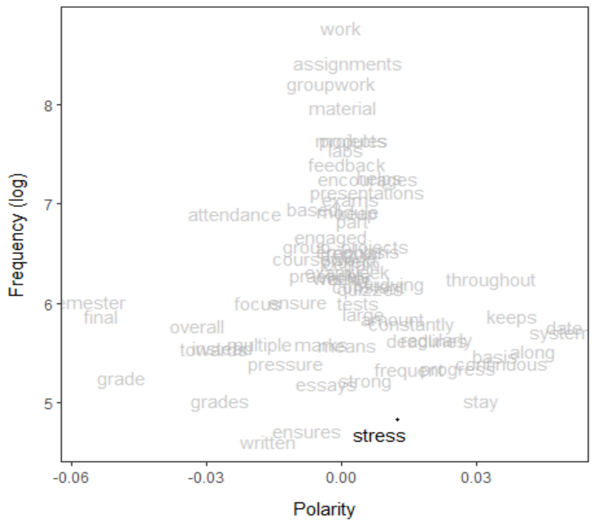


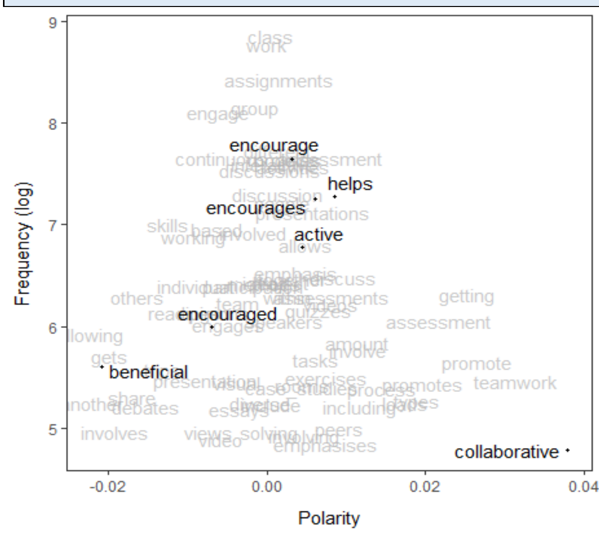
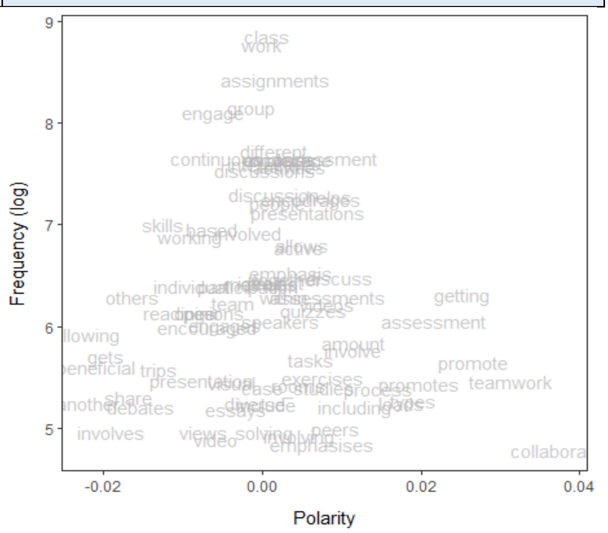
Figure 2.11 presents the words most commonly associated with ‘continuous assessment’ and highlights the words with positive and negative connotations. As can be seen from the charts, the positively associated words for continuous assessment note that it encourages them and ensures their progress. The negative word most often used with reference to continuous assessment is “stress”.

Figure 2.11: Positive and negative words associated with 'Continuous Assessment'

"Continuous Assessment"	
Positive	**Negative**
	

Finally, Figure 2.12 presents the words most commonly associated with 'groupwork' and highlights the words with positive and negative connotations. As can be seen from the charts, the positively associated words for groupwork note that it encourages them and is collaborative and beneficial. There does not appear to be any negative words of note in the second chart.

Figure 2.12: Positive and negative words associated with 'Groupwork'

"Groupwork"	
Positive	**Negative**
	

2.4 Coding Student Responses and Sentiment Analysis

The final step taken in this chapter is to extend the sentiment analysis beyond individual words and instead evaluate whole sentences. Within the survey, students were asked the question, "What does your institution do best to engage students in learning?" As we have seen already students were free to write as much, or as little, as they wanted in answering the question.

In addition, while the question asked could be seen as a direct question, students approached answering it in a myriad of ways. For example, student A could answer the question by saying:

> "I think that the lectures and tutorials provided my institution are the best for engaging students."

Alternatively, student B could say:

> "In the institution the lecturers know my name, and this makes me feel I am more than just a number. I think this helps me engage in learning".

Both are equally valid responses but demonstrate different approaches to answering the question. The first providing a more tangible institutional perspective, the second a more personal and experiential answer.

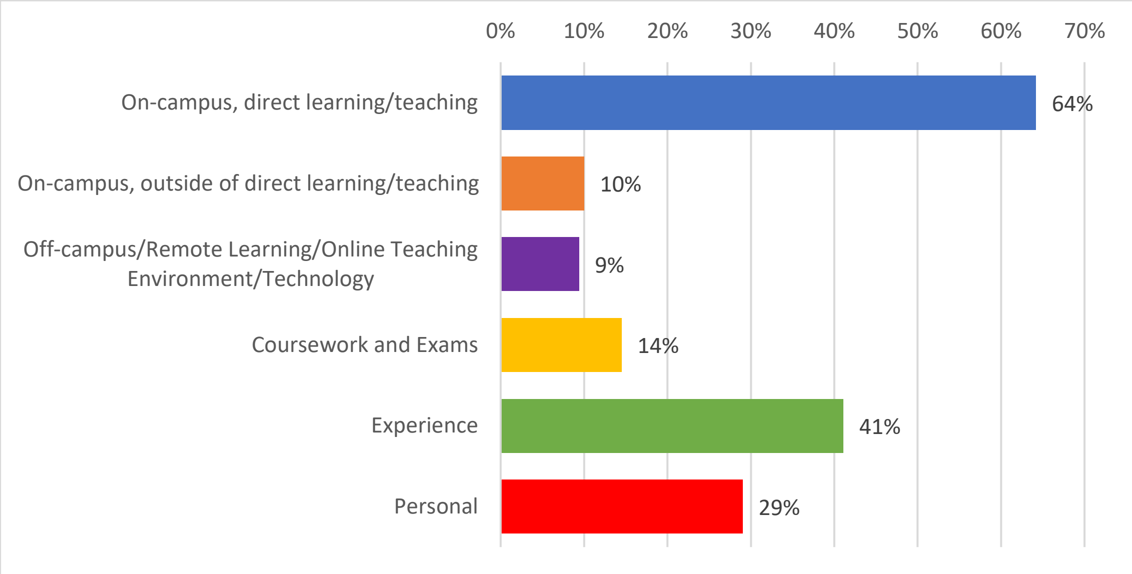
While conducting the preliminary review of students' comments we noted themes and concurrent keywords associated with these themes which were repeated throughout the corpus. This provided us with an initial coding framework. This framework was then supplemented by the analysis reported so far. The coding framework provided in Table 2.3 categorises how students approach answering the question and it is purposely designed along a continuum, with one end being the institutional, and the other personal while simultaneously capturing general versus specific sentiments.

Table 2.3: Coding framework applied to the Q1 corpus

Institutional	
Theme	**Associated Keywords**
On-campus, direct learning/teaching	Tutorials; Workshops; Lectures; Seminars; Labs; Small class/es; Guest speakers; Lecturers; Academic staff; Attendance
On-campus, outside of direct learning/teaching	Academic Learning Centre; Writing centre; Peer-based Learning *(including CÉIM)*; Library; Students' Union; Clubs and Societies; Sport; Facilities
Off-campus/Remote Learning/Online Teaching Environment/ Technology	SULIS/Moodle/Blackboard/Panopto/Brightspace -> *All coded together as Online Teaching Environments*; Clicker; Email; Quizzes
Coursework and Exams	Feedback; Essays; Exams; Dissertation/Thesis; Ask questions; Continuous Assessment
Personal and Experiential	
Theme	**Associated Keywords**
Experience	Collaborative; Discussions; Work placement; Teaching placement; Work experience; Practical elements; Working together; Active participation; Interactive/Interaction; Communication; Hands-on; Help/Support/Assistance; Real life/Real world examples; Problem based learning; Case studies
Personal	Respect; Encourage/Engage; Treated like an adult/equal; "Know my/your name"; Fun; Interesting; Friendly; Approachable; Enjoyable

Within the statistical software, we designed an algorithm that searches for these each of these keywords and iterations thereof, and records where they occur. Of the almost 98,000 comments recorded by students there are over 64,000 cases where at least one of the keywords are mentioned. The analysis presented in the remainder of this section is based on this subset of students[9].

Figure 2.13: Distribution of themes mentioned by students (N = 64,199)



Of the 64,000 cases where at least one of the keywords was mentioned, 64 percent of them are related to on-campus keywords. After this with 41 percent of comments are keywords associated with experience within their HEI. This is closely followed by personal keywords with 29 percent of students' comments. Note that students could mention more than one keyword in their responses and cover a number of themes thus the sum of the percentages sum to more than 100 percent.

[9] Approximately two-thirds of comments contained at least one of the keywords provided in Table 2.3. This is a very good return for such a constrained set of text. Theoretically, it would be possible to code the entire corpus so that all comments fall into at least one category, but it is not particularly practical and beyond a certain point is of limited utility. However, the scope to extend the coding beyond that already outlined to other areas of interest is available to researchers through the R code provided separately.

As has been mentioned already, examining the frequency of themes and how often keywords are mentioned could provide us with a false impression of the sentiment associated with the either. The question asked of students is ‘What does your institution do best to engage students in learning?’, and we may look at the number of students who say ‘lectures’ in their responses as a proxy for what these students think their institutions do best. However, we must be careful here not to divorce keywords of interest from contextual information which provides nuance to the response.

For example, student C may have provided the following response:

> “I love tutorials. They give me a chance to share my reading and discuss with my classmates. They are the complete opposite to my experience with lectures. I find that being handed information in a lecture is a really dry and boring approach to education.”

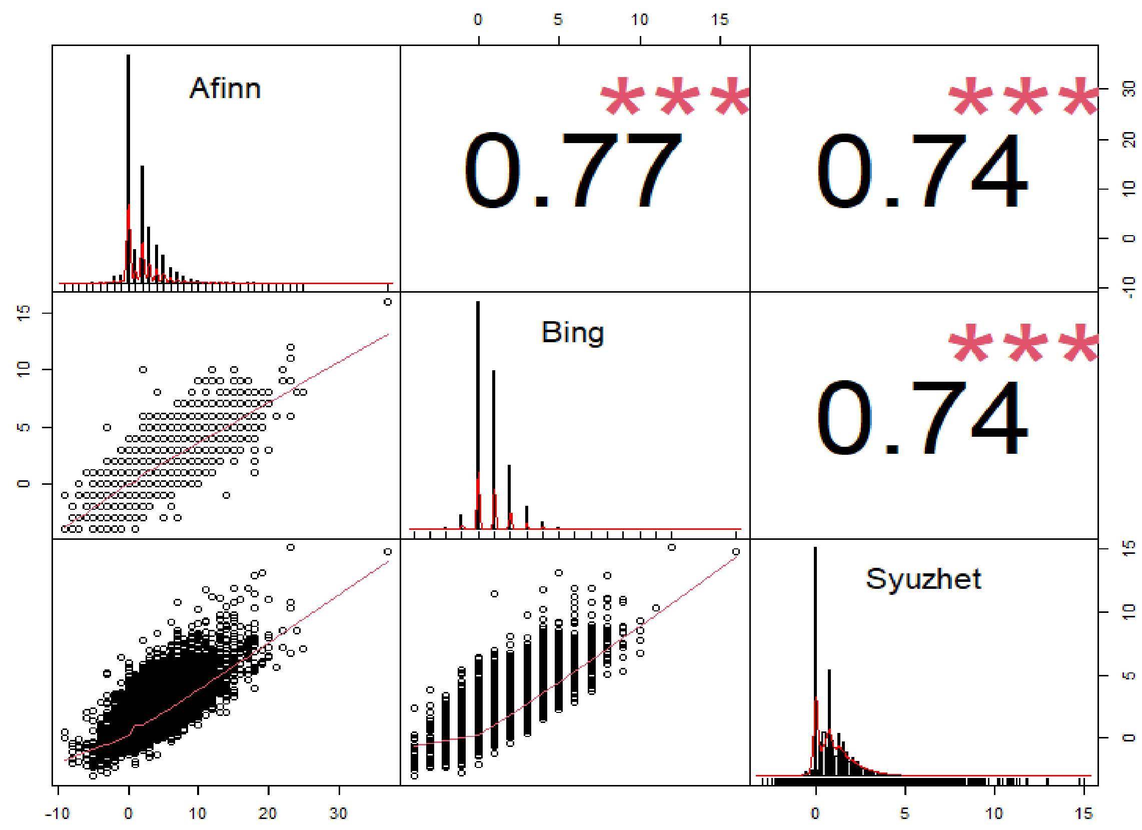
For this student, tutorials are evaluated positively and lectures negatively, but from a naïve frequency perspective, without any further information, as lectures are mentioned twice and tutorials only once, we may incorrectly assume that lectures are viewed more positively because they are mentioned more than tutorials.

While this is easy to distinguish from a single comment, when we are faced with a corpus of the magnitude provided by this question it becomes very appealing to use a frequency heuristic as is it time and resource efficient. We should also be cautious in handling the data not to go the other way and view everything that a student writes with scepticism because at some point we have to view the factors mentioned by students as subjectively important to them, otherwise they would not have mentioned them. As such, the analysis conducted in this section balances these perspectives through combining a frequency-based approach in tandem with sentiment analysis.

The sentiment analysis conducted in this section differs from that of the previous section by extending the analysis beyond words to whole sentences. Sentiments can be broadly classified as positive, neutral or negative. They can also be represented on a numeric scale, to better express the degree to which a body of text contains positive or negative sentiment. The lexicons utilised here evaluate the words present in a students' comment and attach a score to each comment based on the positive, negative and neutral content present in the comment.

The analysis below uses the Syuzhet package for generating sentiment scores, and within this package we have three dictionaries available to us called Afinn, Bing and Syuzhet[10]. Each one was run over the corpus individually as a preliminary test to see if one outperformed the rest.

Figure 2.14: Correlations between Afinn, Bing and Syuzhet sentiment dictionaries



[10] Jockers, M (2020). Syuzhet Package Documentation.
Available from: https://cran.rstudio.com/web/packages/syuzhet/syuzhet.pdf

In Figure 2.14, the distribution of each dictionary variable is shown on the diagonal. On the bottom of the diagonal, the bivariate scatter plots of students' sentiment scores with a fitted line are displayed. On the top of the diagonal, the values of the correlation coefficients are provided along with an indication of the significance level as stars where * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

As can be seen the correlation coefficients between each dictionary are very high and statistically significant, and they all appear to perform in a similar manner. As such, we have presented the Syuzhet results purely because its scores are non-discrete thus have greater potential variation than scores with discrete positions on an axis (for example 0.0, 0.5, 1.0, 1.5, and so on).

At this point, we have run our initial algorithm which records the presence or absence of each of our keywords as a set of indicator variables, and have passed the corpus of students' comments through a sentiment dictionary which has given each comment a sentiment score.

The next step is parsing the keywords with the sentiment analysis and calculating the general sentiment score associated with each keyword. These are presented in the tables below first as a set of summary statistics, with the mean being the one most discussed as this is the best indication of the general sentiment associated with the keyword.

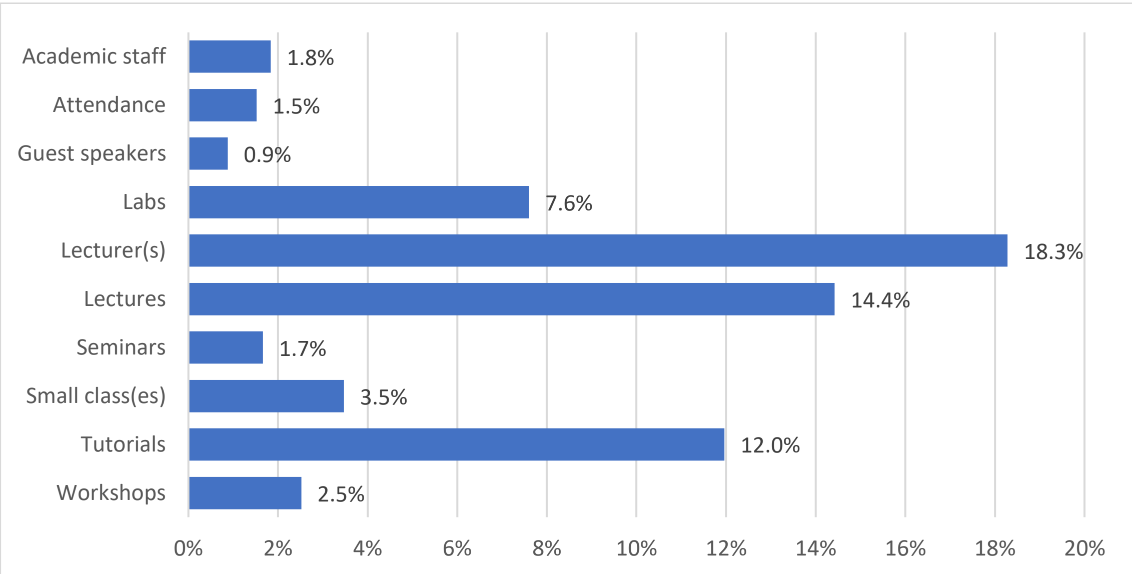
The scale of the sentiment scores is simple to interpret in that higher scores indicate higher levels of positive sentiment, and lower scores indicate higher levels of negative sentiment. This scale also allows for the comparison of scores across keywords, so it is evident below which keywords have significantly higher sentiment scores and lower sentiment scores than others.

While it is important to see how students are evaluating various aspects of their courses, we can also test if different student groups evaluate their courses in the same way. The charts below present the average scores for each keyword across student status and allow us to see if, for example, first year undergraduates have a more positive sentiment score for lectures than taught undergraduates. The mean sentiment scores across student groups have been presented with 95% confidence intervals around them. Where these do not overlap with one another, we can say that the mean sentiment scores across groups and across keywords are significantly different from one another.

It is worth highlighting that this approach is not without its imperfections, but there is an inherent trade-off between human coding, time, and error. In general, from examining the scores associated with each sentence in the corpus, the sentiment dictionaries do a good job of capturing the sentiment contained within comments even if they lack the ability to parse nuance, irony and sarcasm. It is more than this methodology excels at evaluating the underlying sentiment when the wealth of material is too much for a human coder and would fall prey to biases in human coders.

The greatest strength of computer-assisted content analysis is that all material is treated equally. In contrast for a human coder there is just too much material; as noted elsewhere, the corpus of Q1 runs to 4000 A4 pages of text, for which it is estimated that at an average reading speed would take a human coder 110 hours to read through once. This amount of effort and concentration would bring with it the risk of error due to fatigue, recency bias, and so on. Whereas our software evaluates the corpus in seconds.

Figure 2.15: Distribution of 'on-campus, direct learning/teaching' keywords (N = 41,193)



To move onto the sentiment analysis itself. As pointed out in Figure 2.13, keywords associated with being on-campus and having direct interaction with teaching and learning are present in 64 percent of the subsetted corpus. Within this theme, the frequency of its component keywords, are presented in Figure 2.15. From this chart, one can see that lecturers are most often mentioned in students' comments, closely followed by lectures, and then tutorials. The other keywords are mentioned much less often with labs being mentioned in 7.6 percent of comments down to guest speakers who are only mentioned in less than one percent of comments.

The sentiment associated with each keyword does not follow the pattern demonstrated by the frequency of keywords, illustrating the point made above that by itself, frequency is not a good indicator of what students think helps them engage in learning at their institution. While being only mentioned in close to two percent of students' comments, academic staff have the highest mean sentiment score in comments where the keyword is used at 2.18.

Table 2.4: Summary statistics for 'on-campus, direct learning/teaching'

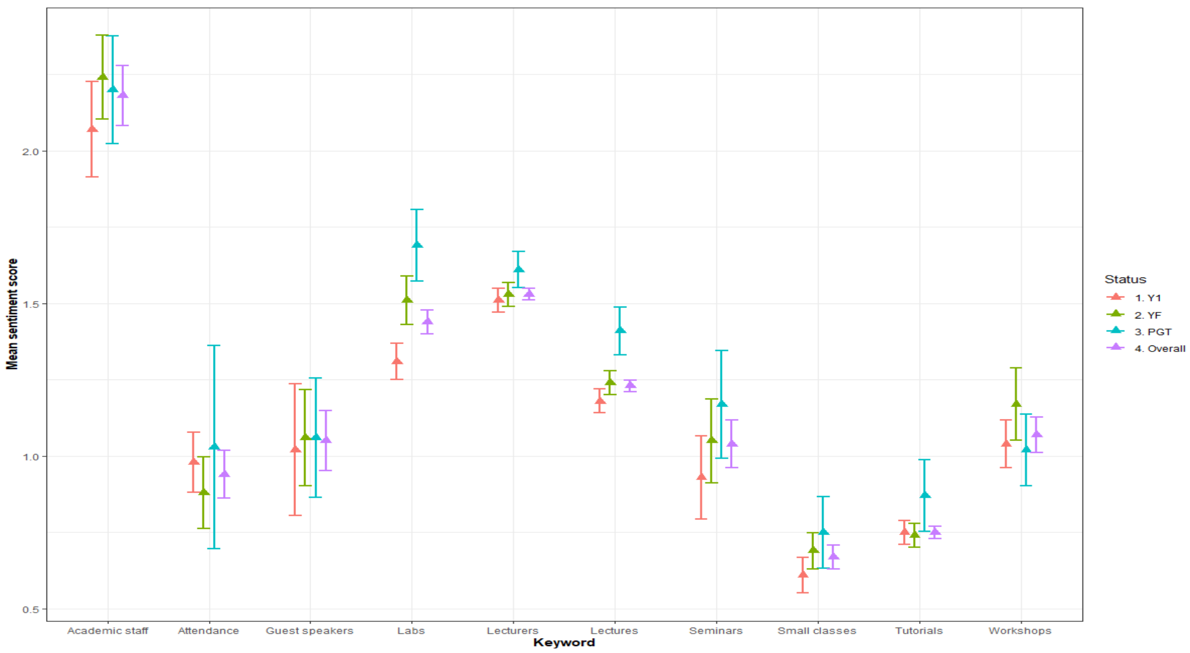
	Minimum	**Mean**	Median	Maximum	Standard deviation	Count
Academic staff	-0.60	**2.18**	1.90	14.75	1.55	1,179
Attendance	-2.05	**0.94**	0.40	11.00	1.19	977
Guest speakers	-0.75	**1.05**	0.75	7.90	1.30	563
Labs	-2.00	**1.44**	1.15	11.85	1.44	4,882
Lecturer(s)	-2.75	**1.53**	1.25	14.75	1.38	11,738
Lectures	-2.25	**1.23**	0.80	13.05	1.34	9,259
Seminars	-0.90	**1.04**	0.60	11.85	1.42	1,068
Small classes	-1.35	**0.67**	0.00	10.20	1.05	2,229
Tutorials	-3.00	**0.75**	0.25	10.80	1.12	7,681
Workshops	-1.00	**1.07**	0.80	9.70	1.30	1,617

While being present in more students' comments, lecturers, lectures, and tutorials have mean sentiment scores of 1.53, 1.23, and 0.75 respectively. This indicates that comments that contain these keywords are a combination of positive and negative sentiments. The keyword attendance is typically combined with 'compulsory' in the comments and does not appear to be evaluated positively by students with a mean score of 0.94. Finally, contrary to expectations, small classes do not appear to positively evaluated by students, with an average sentiment score of only 0.67.

Figure 2.16 presents the average sentiment scores for students along with the average sentiment scores for each keyword contained in the 'on-campus, direct teaching/learning' theme for first year undergraduates (Y1), final year undergraduates (YF), and taught postgraduates (PGT). The sentiment scores are also complimented with 95% confidence intervals which where these do not overlap, show quickly and in a readily interpretable format, statistically significant differences across groups and across keywords.

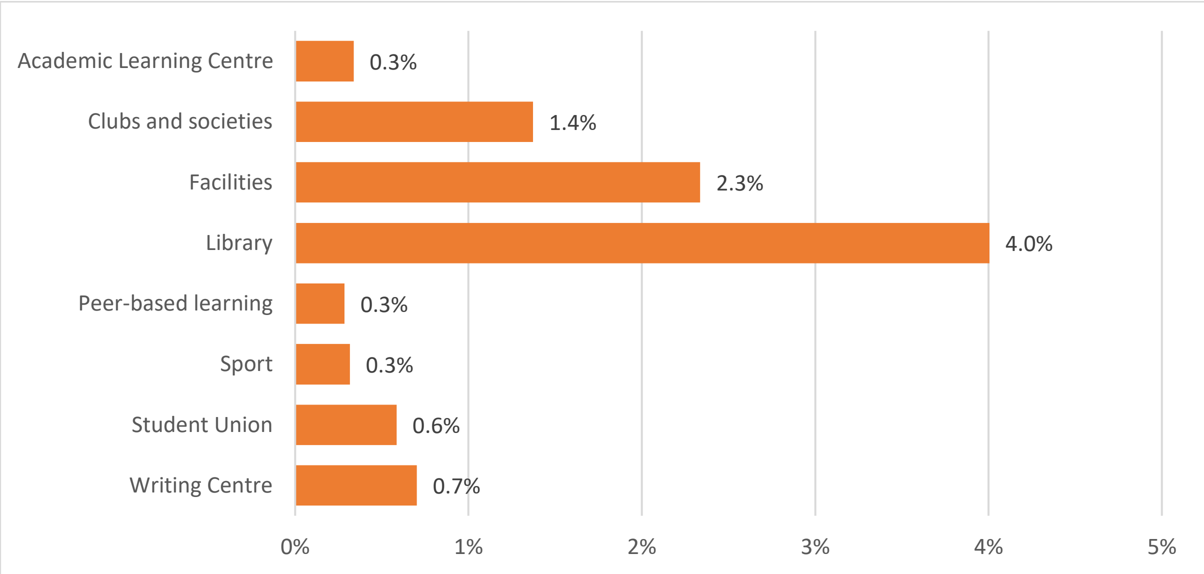
As noted above, comments that mention academic staff have higher average sentiment scores than all other keywords in this theme, though these are not significant across student groups. Some differences across groups appear for labs, lectures and lectures, with taught postgraduates providing these keywords with higher average sentiment scores than first year undergraduates.

Figure 2.16: Average sentiment scores of 'on-campus, direct learning/teaching' keywords (with 95% confidence intervals) across student status



Keywords associated with being on-campus but indirectly interacting with teaching and learning were present in only 10 percent of the subsetted corpus (cf. Figure 2.13). Within this theme, the frequency of its component keywords, are presented in Figure 2.17. From this chart, one can see libraries are most often mentioned in students' comments, closely followed by the general facilities provided by their HEI, and then the clubs and societies of their HEI.

Figure 2.17:Distribution of 'on-campus, outside of direct learning/teaching' keywords (N=6,418)



The common thread to each of the keywords contained in this theme is that their presence or absence helps students to help themselves, are facilities provided by institutions to assist students in learning how to learn outside of their course, or are factors that assist students to balance academic life with the rest of student life. Interaction with each of these keywords is largely the decision of individual students not their institutions.

Table 2.5: Summary statistics for ‘on-campus, outside of direct learning/teaching’ keywords

	Minimum	**Mean**	Median	Maximum	Standard Deviation	Count
Academic Learning Centre	-2.50	**1.45**	1.20	6.80	1.08	217
Clubs and societies	-0.50	**1.30**	0.75	13.05	1.67	881
Facilities	-1.85	**1.90**	1.60	13.05	1.48	1,500
Library	-1.85	**2.10**	1.75	11.80	1.44	2,571
Peer-based learning	-0.95	**0.59**	0.05	4.35	0.89	183
Sport	-1.10	**1.30**	0.75	11.85	1.76	203
Student Union	-1.25	**1.90**	1.55	11.90	1.68	376
Writing Centre	-0.30	**1.79**	1.50	5.15	1.46	451

From Table 2.5 one can see that the libraries of each institution have the highest mean sentiment score in comments where the keyword is used at 2.10, closely followed by general facilities provided by institutions and Students’ Unions both with an average sentiment score of 1.90.

Figure 2.18 presents the average sentiment scores for students along with the average sentiment scores for each keyword contained in the ‘on-campus, outside of direct teaching/learning’ theme for first year undergraduates (Y1), final year undergraduates (YF), and taught postgraduates (PGT). The sentiment scores are also complimented with 95% confidence intervals which where these do not overlap, show quickly and in a readily interpretable format, statistically significant differences across groups and across keywords. Some categories have larger confidence intervals around the mean sentiment score because the lower number of mentions has an influence on the standard error used to construct the confidence intervals. Given this, there does not appear to be any significant differences between student groups, but the mean sentiment scores for libraries, facilities and clubs and societies appear to be significantly more positive than the mean sentiment score for peer-based learning.

Figure 2.18: Average sentiment scores of ‘on-campus, outside of direct learning/teaching’ keywords (with 95% confidence intervals) across student status

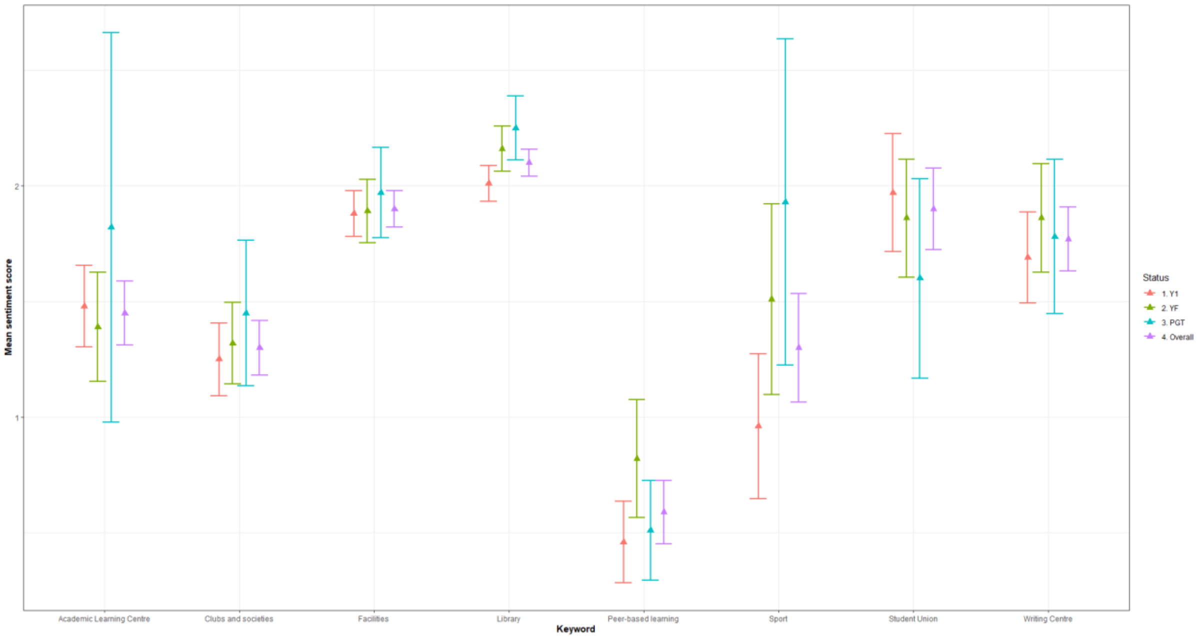
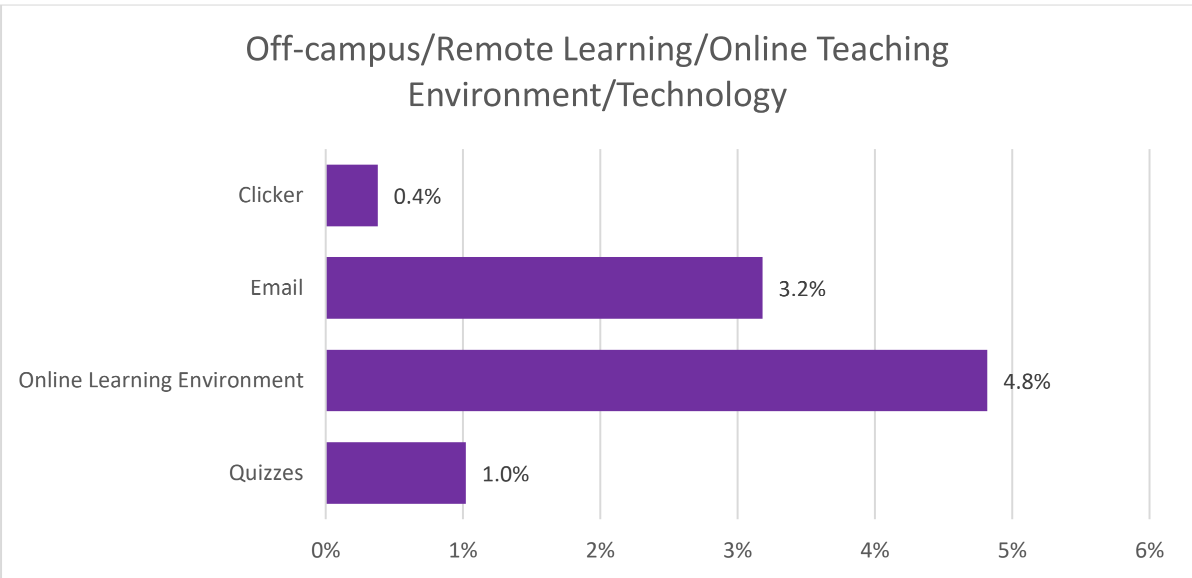


Figure 2.19: Distribution of ‘off-campus/remote learning/online teaching environment/technology/coursework and exams’ keywords (N=6,033)



Keywords associated with the technology that assisted them to work off-campus (online learning environments/email) or interacted with on-campus (clicker or quizzes in lectures) were present in only 9 percent of the subsetted corpus (cf. Figure 2.13). Within this theme, the frequency of its component keywords, are presented in Figure 2.19. From this chart, one can see the online learning environment is most often mentioned in students' comments. This category was constructed from the individual and varied platforms such as Blackboard, Moodle and so on which are using across HEIs.

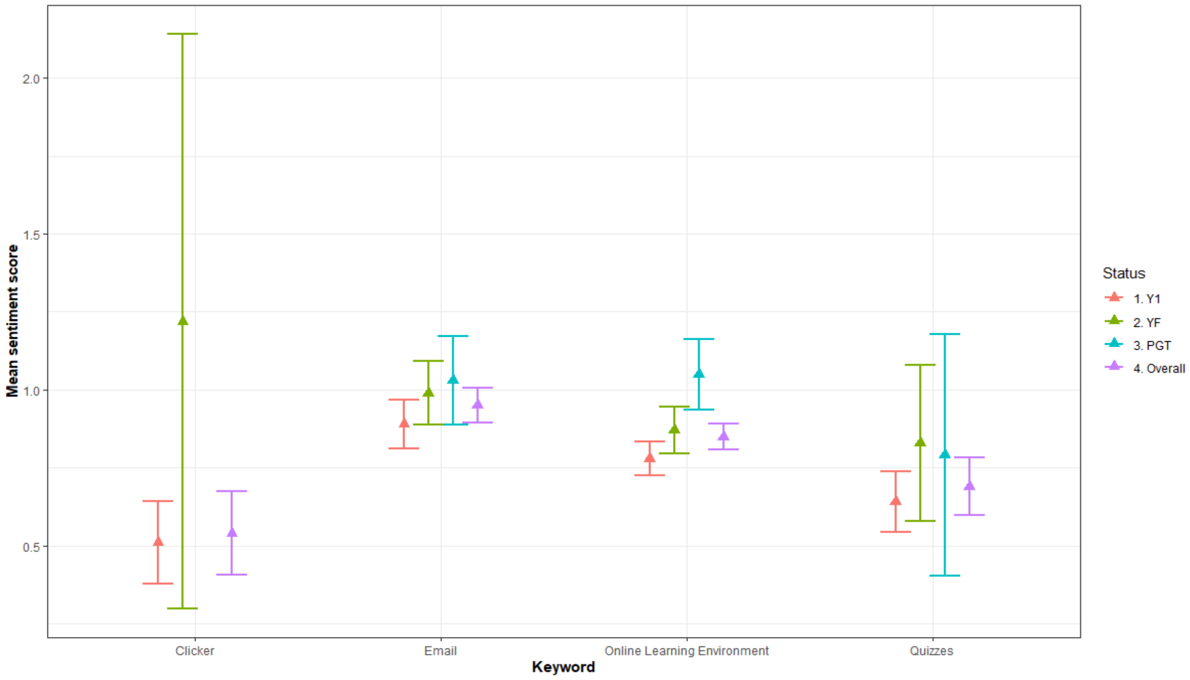
Table 2.6: Summary statistics for 'off-campus/remote learning/online teaching environment/technology/coursework and exams' keywords

	Minimum	**Mean**	Median	Maximum	Standard Deviation	Count
Clicker	-1.00	**0.54**	0.00	7.50	1.07	243
Email	-1.35	**0.95**	0.50	10.45	1.32	2,042
Online Learning Environment	-1.45	**0.85**	0.50	11.85	1.16	3,093
Quizzes	-1.25	**0.69**	0.00	8.40	1.20	655

From Table 2.6 one can see that these keywords have in general lower sentiment scores than categories in themes already discussed, indicating that students may not have many positive interactions with these facets of higher education or a lower general opinion about them with regard to how much they engage students.

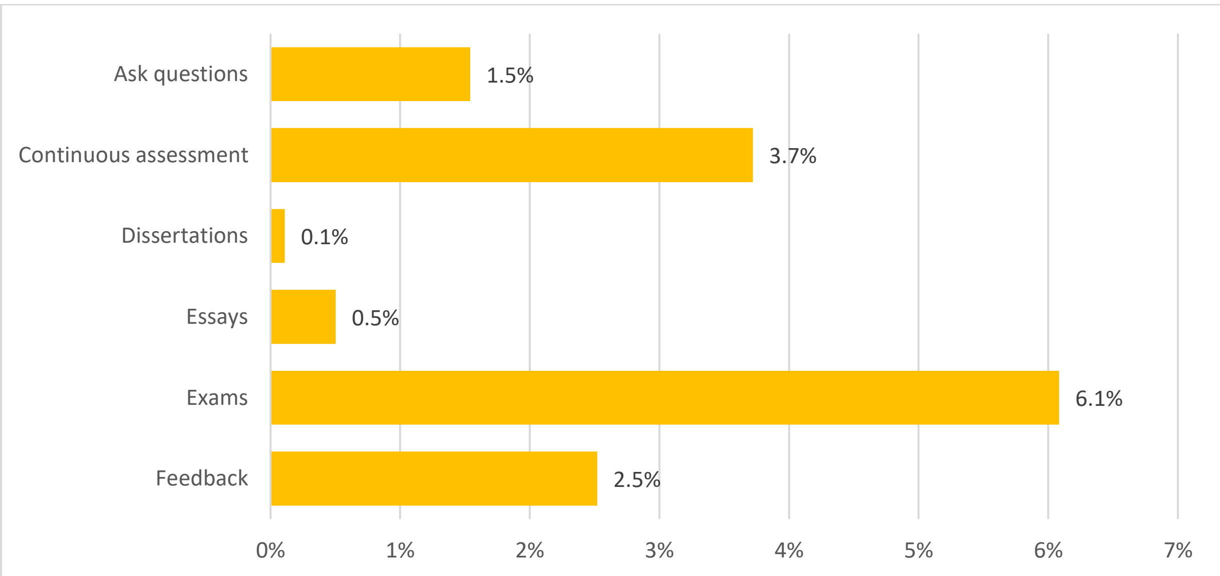
A similar pattern is present in Figure 2.20 presents the average sentiment scores for students along with the average sentiment scores for each keyword contained in the 'on-campus, outside of direct teaching/learning' theme for first year undergraduates (Y1), final year undergraduates (YF), and taught postgraduates (PGT). Note that there are a low number of mentions of 'clicker' for final year students and none at all for taught postgraduates. Thus, we see a large confidence interval and no confidence interval respectively for each group for this keyword. Aside from this, taught postgraduates appear to have a significantly higher mean sentiment score for the various online learning environments provided by their HEIs than first year undergraduates.

Figure 2.20: Average sentiment scores of 'off-campus/remote learning/online teaching environment/technology' keywords (with 95% confidence intervals) across student status



Keywords associated with coursework and exams were present in only 14 percent of the subsetted corpus (cf. Figure 2.13). Within this theme, the frequency of its component keywords, are presented in Figure 2.21. From this chart, one can see exams are most often mentioned in students' comments, closely followed by continuous assessment, and receiving feedback.

Figure 2.21: Distribution of 'coursework and exams' keywords (N=9,297)

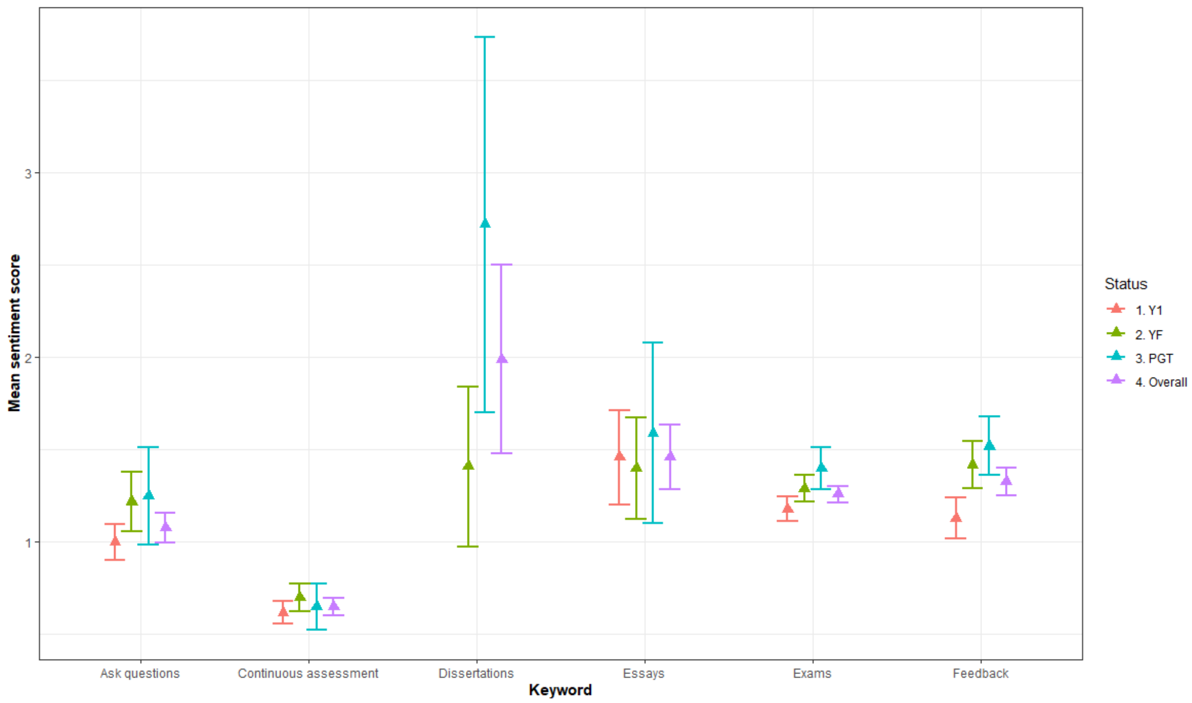


Within this theme, dissertations have the highest average sentiment score with 1.99 though this keyword was only mentioned 71 times in the corpus we should be careful of placing too much emphasis on this. Looking at Table 2.7 and Figure 2.22 together shows that exams, essay receiving feedback and the ability to ask questions are all similarly evaluated by students. In contrast, continuous assessment has a significantly lower average sentiment score.

Table 2.7: Summary statistics for 'coursework and exams' keywords

	Minimum	**Mean**	Median	Maximum	Standard Deviation	Count
Ask questions	-1.50	**1.08**	0.80	8.65	1.29	989
Continuous assessment	-1.50	**0.65**	0.00	9.70	1.12	2,390
Dissertations	-1.25	**1.99**	1.75	11.80	2.20	71
Essays	-1.00	**1.46**	1.00	9.70	1.62	323
Exams	-2.15	**1.26**	1.00	12.90	1.43	3,906
Feedback	-1.45	**1.33**	0.90	11.05	1.54	1,618

Figure 2.22: Average sentiment scores of 'coursework and exams' keywords (with 95% confidence intervals) across student status



Keywords associated with experiential factors associated with learning contributed to 41 percent of comments (cf. Figure 2.13). However, Figure 2.23 shows that a large component of this related to receiving help, support, or assistance from their institution.

Figure 2.23: Distribution of 'experience' keywords (N=26,371)

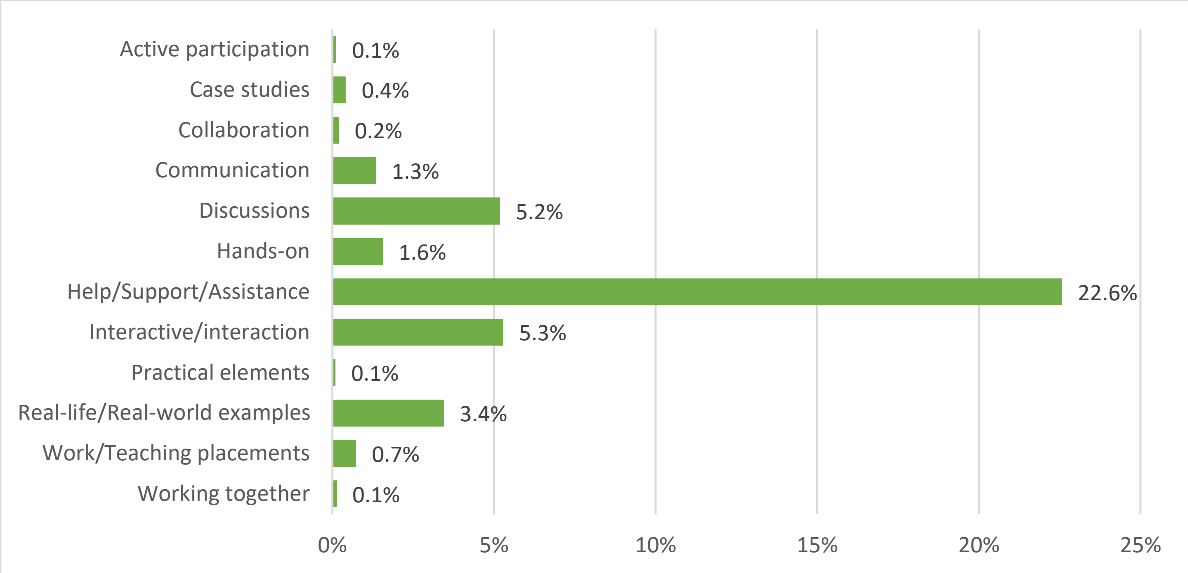


Table 2.8 shows that receiving help/support/assistance from the HEI received a high average sentiment score potentially indicating that, when needed, institutions are able to provide this for students.

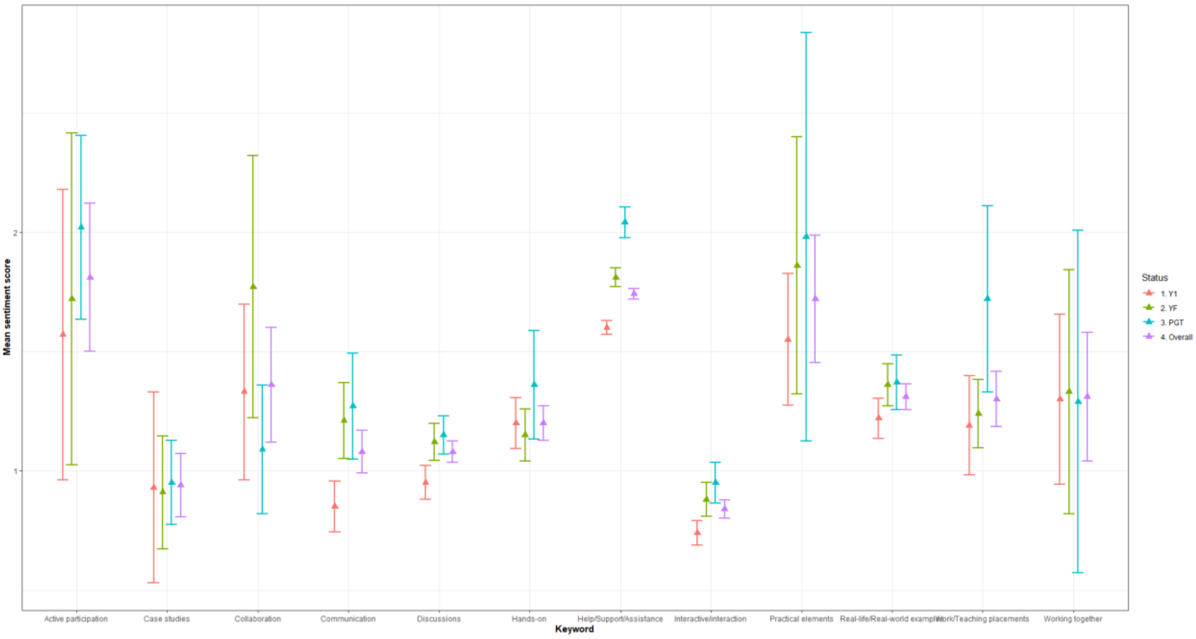
Table 2.8: Summary statistics for 'experience' keywords

	Minimum	**Mean**	Median	Maximum	Standard Deviation	Count
Active participation	0.35	**1.81**	1.25	6.45	1.38	76
Case studies	-0.25	**0.94**	0.75	6.50	1.11	267
Collaboration	0.00	**1.36**	0.85	9.40	1.40	130
Communication	-1.60	**1.08**	0.75	14.75	1.35	864
Discussions	-1.75	**1.08**	0.65	10.45	1.30	3,327
Hands-on	-0.60	**1.20**	1.00	8.30	1.17	1,000
Help/Support/Assistance	-2.00	**1.74**	1.45	15.05	1.40	14,488
Interactive/Interaction	-1.75	**0.84**	0.60	9.05	1.13	3393
Practical elements	1.00	**1.72**	1.25	5.60	1.04	58
Real-life/Real-world examples	-2.50	**1.31**	1.05	12.90	1.30	2,213
Work/Teaching placements	-2.05	**1.30**	1.00	7.90	1.29	472
Working together	0.05	**1.31**	1.05	6.1	1.25	83

Figure 2.24 presents the average sentiment scores for students along with the average sentiment scores for each keyword contained in the 'experience' theme for first year undergraduates (Y1), final year undergraduates (YF), and taught postgraduates (PGT). Some categories have larger confidence intervals around the mean sentiment score because the lower number of mentions has an influence on the standard error used to construct the confidence intervals.

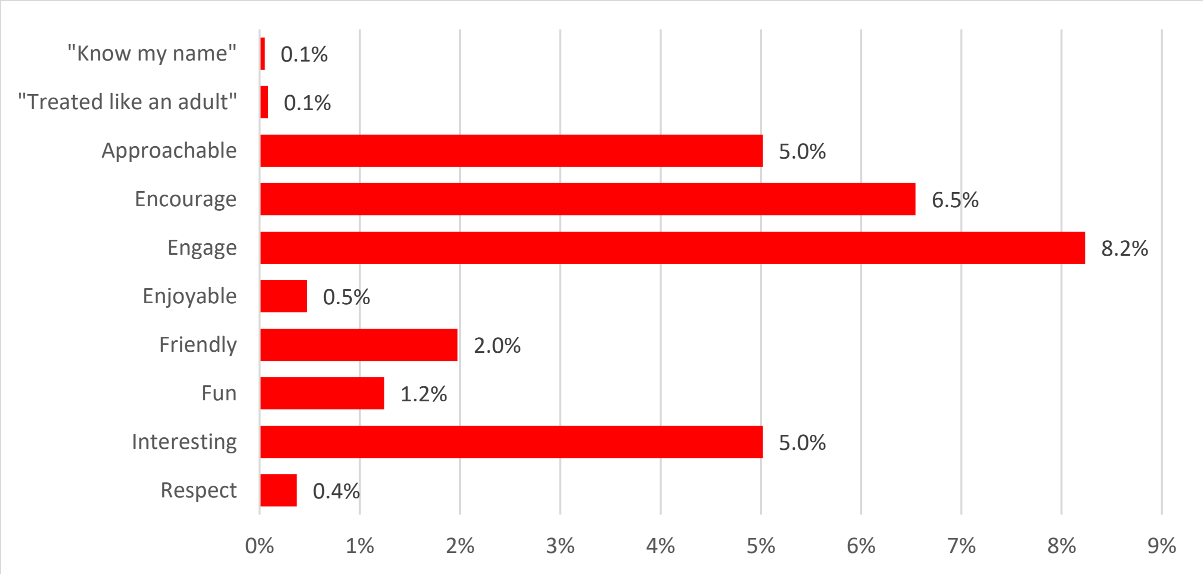
Despite this, there appears to be a number of significant differences between student groups with first year undergraduates having lower average sentiment scores for help/support/assistance, discussions, and communication than taught postgraduates.

Figure 2.24: Average sentiment scores of 'experience' keywords (with 95% confidence intervals) across student status



Finally, Keywords associated with students' personal interaction and relationship with their HEI formed 29 percent of the subsetted corpus (cf. Figure 2.13). Within this theme, the frequency of its component keywords, are presented in Figure 2.25. From this chart, one can see that being engaged is most often mentioned, along with being encouraged, and finding things interesting.

Figure 2.25: Distribution of 'personal' keywords (N=18,625)



With regard to the average sentiment scores for these keywords, a case can be made to err on the side of caution here, because of the connotations present in each word. It is highly likely that the average sentiment scores for each of these keywords are higher than they should be because the words themselves have positive associations. As a result, their presence in a sentence biases the overall score towards being more positive than it would be if the keyword was not present.

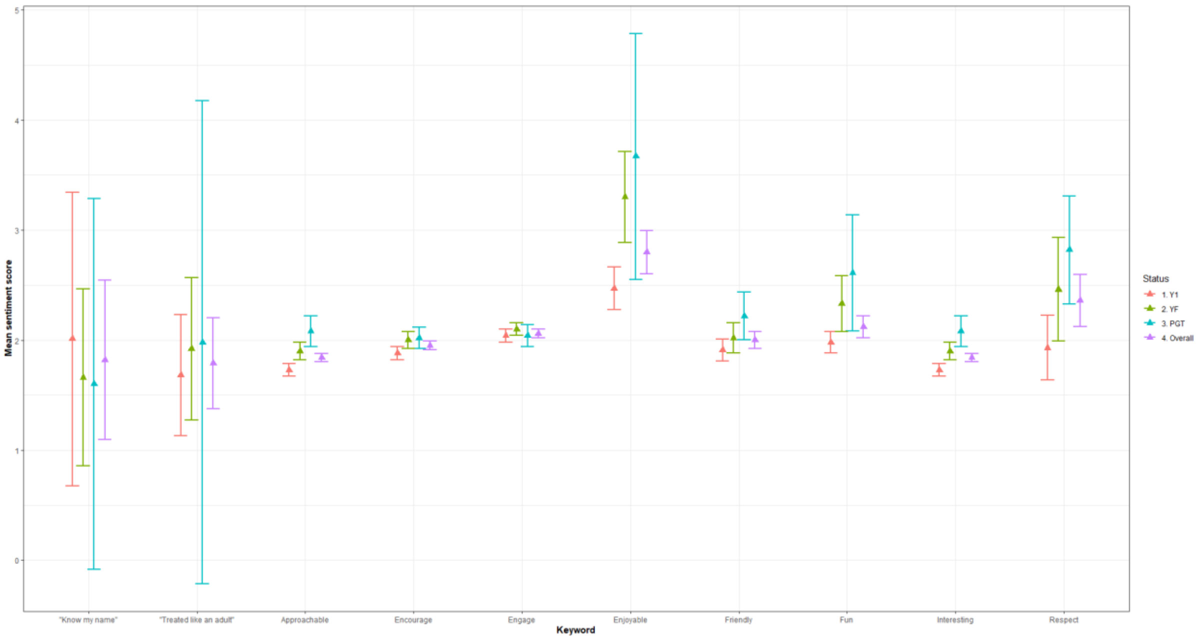
Table 2.9: Summary statistics for 'personal' keywords

	Minimum	**Mean**	Median	Maximum	Standard Deviation	Count
"Know my name"	0.00	1.82	1.35	10.80	2.14	33
"Treated like an adult"	-0.25	1.79	1.60	5.40	1.53	53
Approachable	-1.15	1.84	1.55	15.05	1.31	3,222
Encourage	-2.15	1.95	1.55	11.90	1.38	4,201
Engage	-1.75	2.06	1.75	11.90	1.52	5,286
Enjoyable	0.00	2.80	2.40	14.75	1.79	304
Friendly	-1.25	2.00	1.55	9.25	1.36	1,268
Fun	-1.85	2.12	1.75	13.05	1.49	797
Interesting	-1.15	1.84	1.55	15.05	1.31	3,222
Respect	-1.95	2.36	2.05	15.05	1.91	239

This of course is not to say that these factors are not important to students, as has already been noted, the fact that these keywords have been mentioned (outside of 'engage' which they are primed to use due to its presence in the original question) can be taken as some measure of the importance of these for students, regardless of whether their institution actually meets them.

Despite all of this, it is worth noting how similar all the average sentiment scores - across keywords and across groups - are to one another. Figure 2.26 illustrates this best, with the mean score for most of the keywords along with confidence intervals oscillating around the 2.0 mark on the chart.

Figure 2.26: <u>Average sentiment scores of 'personal' keywords (with 95% confidence intervals) across student status</u>



2.5 Conclusions

This chapter has broken down the analysis undertaken to address the question posed to students; “What does your institution do best to engage students in learning?” The analysis conducted has demonstrated that students answered this question in a myriad of ways, but from this almost overwhelming amount of material certain patterns emerged. From breaking down students’ comments into keywords, and then examining the general sentiment present in sentences that contained these keywords, we can provide some answers to the question asked of students.

Table 2.10: Ranking of keywords with highest average sentiment scores

Rank	Keyword	Mean Sentiment Score
1	Academic staff	2.18
2	Library	2.10
3	Dissertations*	1.99
4	Facilities	1.90
5	Student Union	1.90
6	Writing Centre*	1.77
7	Help/Support/Assistance	1.74
8	Practical elements	1.72
9	Lecturer(s)	1.53
10	Essays	1.46
11	Academic Learning Centre*	1.45
12	Labs	1.44
13	Collaboration	1.36
14	Feedback	1.33
15	Real-life/Real-world examples	1.31

In general, it appears that students think that the academic staff of their institution are the best for engaging students in learning. This is followed by the library services, and general facilities provided by their HEI. The support provided for dissertations is also highly placed by is infrequently mentioned in the corpus thus has an asterisk next to it in Table 2.10 to highlight the fact that support for this assertion is weaker. Academic Learning Centres and Writing Centres are also highlighted in this way, as they are positively evaluated by students, they are not present in every HEI. Though from a policy perspective, the fact that they are so positively evaluated by students where they are present would strengthen an argument for their use in HEIs where they are so far absent.

Keywords associated with 'personal' experiences have been left from the table due to the fact that their mean scores may be positively biased due to the positive associations inherent to the keywords themselves. This, however, does not mean that HEIs should ignore these facets of students' experiences. Instead, if greater attention were paid to these facets by HEIs, academic staff and other points of contact between academic institutions and students, one could see a virtuous cycle which reinforces other aspects of students experiences and increases student engagement.

The next chapter moves onto addressing how students evaluated the second qualitative question they were asked; "What could your institution do to improve students' engagement in learning?"

3. What could your institution do to improve students' engagement in learning?

This chapter of the report covers the analysis undertaken to provide us with some answers to the second open-text question posed to students; “What could your institution do to improve students’ engagement in learning?” Which for the shorthand purposes in the rest of this chapter will be referred to as Q2.

In contrast to Q1, which was worded in a manner to evoke a precise response, such as an item, or institutional provision that students regarded as being successful in engaging students, Q2 poses a more hypothetical query, which asks about a change that could be made, and asks students to weigh up what they have received from their HEI and from this, evaluate what change could be made to improve students’ engagement in learning.

The structure of the question is similar to an equation which asks, what change needs to be made to X to see an improvement/increase in Y where Y is students’ engagement in learning. As a result of the way the question is asked, responses tend to point to an item and provide a description of degree, or provide a modifier/intensifier, for example, typical responses to Q2 included statements such as ‘’provide more continuous assessment”, “better lecturers”, “fewer exams”, “longer library opening hours” and so on.

The analysis that this chapter discusses follows a similar format to that used in Q1 but with a few significant deviations to capture this fact that standard responses from students provide **items** that students think should be altered to improve student engagement and some **measure of direction and degree** (more/less/better/fewer etc).

In addition, the sentiment analysis used in the previous chapter is of less utility here because of the prevalence of a small group of words used to indicate direction and degree. This is discussed further below. As such, there is less need to capture what students are feeling about keywords through the words used in their responses because there is:

a. Much less variety in the words used to describe what they would change about these keywords.
b. The words used to describe direction and degree specify through their usage how they would affect items, thus rendering it unnecessary to further examine the sentiment that these words contain.

However, the first task facing the researcher is parsing the material into a comprehensible format. This is again, a daunting task due to the volume of material students have provided. For Q1 students' responses would fill 4,000 A4 pages, whereas for Q2, students' responses would fill 3,850 pages of A4. The next section covers the preliminary meta-analysis of Q2 and describes how the corpus was cleaned for further analysis.

3.1 Meta-analysis and Frequency of Characters

As with Q1, responses to Q2 were open-text and free of any character limits. As such, respondents could provide as much or as little text as they wanted. This freedom has resulted in considerable variation in the amount of text students provided. Table 3.1 provides some summary statistics of the number of characters used by students in their responses. Characters include letters, numbers, punctation marks and spaces. This table shows that on average, the length of a response was almost 90 characters long. The median length was 52 characters. The minimum length was three characters, and the maximum was 9,211, which is about three pages of A4 text.

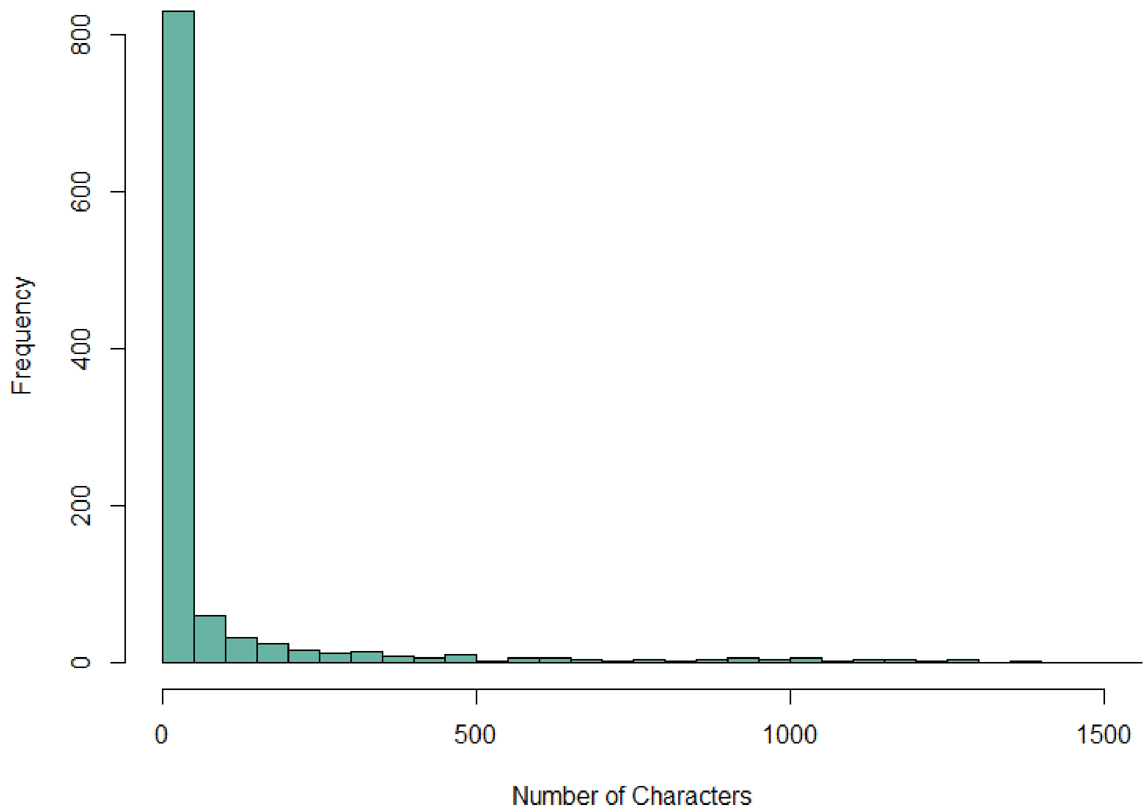
Almost 92,000 students provided responses to Q2 compared against close to 98,000 for Q1. However, the average response to Q2 was longer than those provided for Q1, as shown by the higher mean, median, and maximum.

Table 3.1: Summary statistics of the number of characters in Q2

Mean	Median	Standard deviation	Interquartile Range	Minimum	Maximum	Total N
88	52	129	80	3	9,211	91,192

While respondents provided more information in their responses for Q2, the distance between the mean and the median again indicate a non-normal distribution of characters much like that seen in Figure 2.1 in the previous chapter. A histogram of this distribution is provided in Figure 3.1 and again shows that about fifty percent of responses use around fifty characters, and the rest use more than this. However, it is worth noting that the x-axis in Figure 3.1 has been limited to 2000 characters. There are a few responses outside of this limit but plotting them made the chart difficult to read and interpret.

Figure 3.1: Distribution of the number of characters in students' responses to Q2



The skewed distribution of characters is again, not unexpected, and would only be a concern if different groups of students had different patterns of responding to the question. To ensure that this is not the case, Figure 3.2 presents three boxplots of the distribution of characters in students' responses to Q2 by student status (first year undergraduates (Y1), final year undergraduates (YF), and taught postgraduates (PGT)).

As can be seen in this chart, each group appears to follow the same skewed distribution shown in Figure 3.1. However, first year undergraduates appear to provide on average, shorter responses to Q2 than the other groups as the median for this group shown by the black central bars within the boxplot is marginally lower than the overall median, which is presented as a red dashed line on the graph. Again, this is not wholly unexpected as final year undergraduates and taught postgraduates by virtue of the longer length of time they have spent in third level education have more experience to draw upon when providing an answer, and as such provide more detail in their answers. Though substantively, the relative difference across groups is minor and does not affect the rest of the analysis discussed below.

Figure 3.2: Boxplots of the distribution of characters in responses to Q2 by Student Status

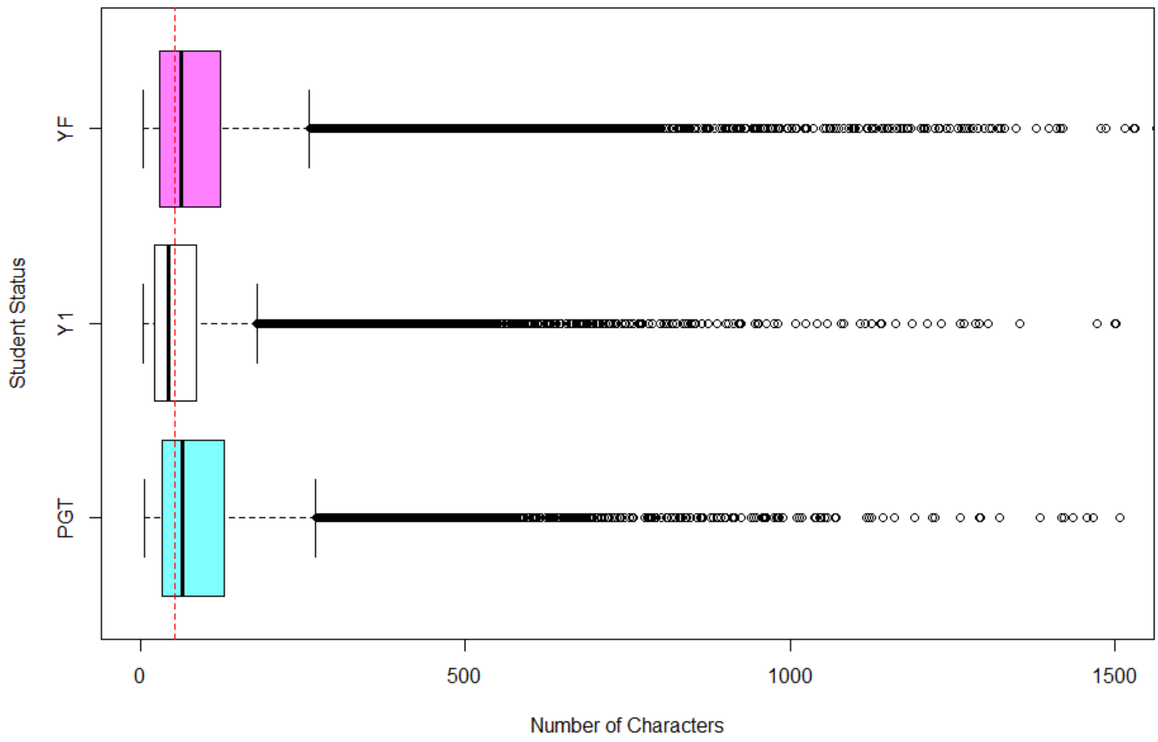
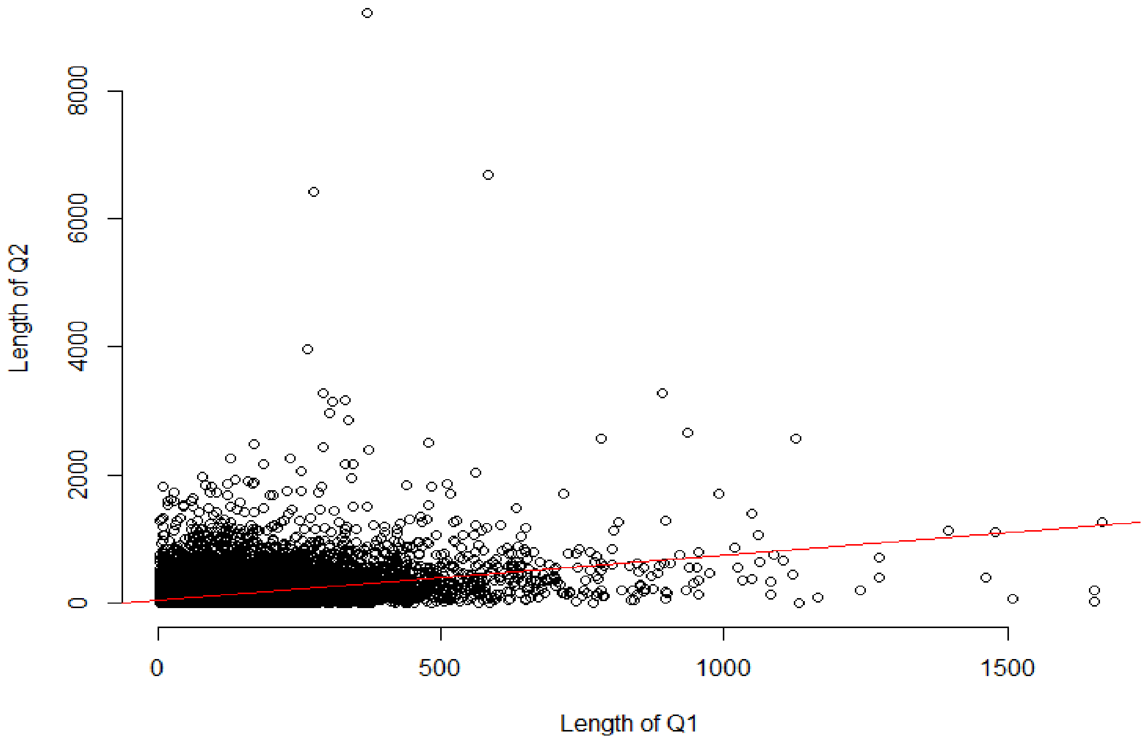


Figure 3.3 shows the lengths of Q1 and Q2 plotted against one another where comments were provided for both Q1 and Q2. As can be seen from the chart, there is a positive correlation between them with short comments for Q1 being associated with short comments for Q2 and longer comments for Q1 being associated with longer comments for Q2. This is shown through the regression line plotted on top of the cases, though this is not a particular strong relationship as the line while positive, is relatively flat. This is supported by the correlation coefficient, calculated separately, which is only 0.43.

Figure 3.3: Correlation between the frequency of characters used in responses to Q1 and Q2



As mentioned in Chapter 1, the number of characters students use to provide a response tells us something about the willingness of students to provide information but does not tell us anything about the information provided. To gain an understanding of the material provided, we must be confident of the quality of the material to begin with, and to ensure this, the data goes through a series of steps to remove any extraneous content which would undermine the analysis. These steps take the corpus of material from a 'raw' to 'clean' state, though throughout these steps a copy of the original data is kept as a reference and quality control check.

The first invaluable check is gaining a familiarity with the data through reading through the material provided. This is time-intensive but provides a basis from which to build an understanding of how students have approached the question asked of them. Because of the volume of material provided by students it remains impractical to read everything but by combining the immersion in the text with the spellcheck component it is still possible to get a solid grounding in recurring topics, keywords and approaches. Text provided in Irish was translated into English, comments below a certain length (fewer than three characters) were removed, and the same compounding of common multiphrase words (continuous assessment, academic staff, and so on) which should be examined by the statistical software together was also conducted[11].

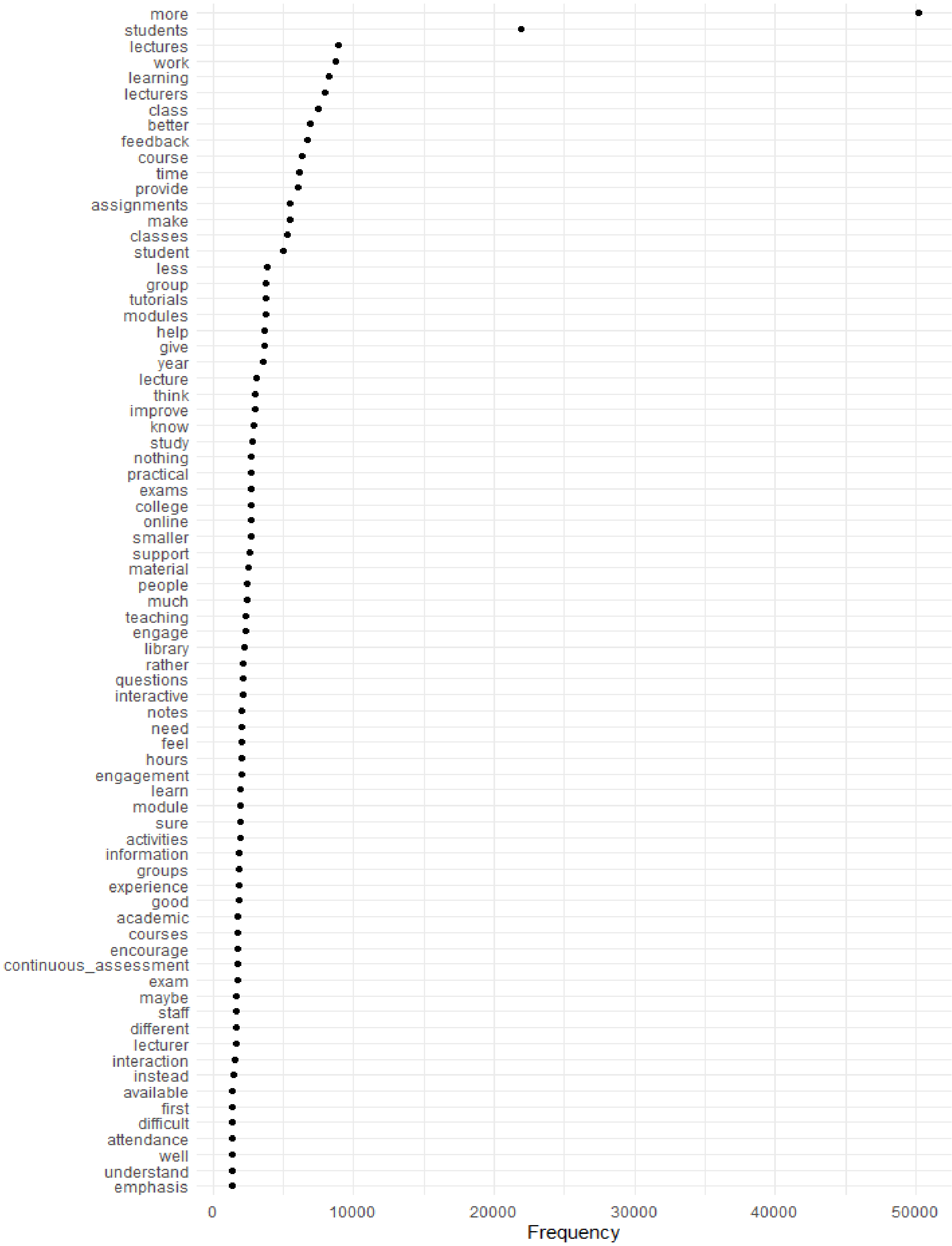
Once these steps have been completed the data is ready for entry into the statistical software, so that punctuation and symbols can be removed, along with common stop words. Appendix A provided a comprehensive list of the stop words removed by the Quanteda package in R. However, for this portion of the project the word ‘more’ was added back into the list of allowed words rather than being removed, as this word was invaluable in the analysis below, as it is the most common way students had of articulating the changes they wished to see in their HEIs to improve student engagement.

Figure 3.4: Word cloud of the 75 most frequently used words

The remaining corpus was then ‘tokenised’ based on the whitespace between words so the analysis could begin in earnest. Figure 3.4 presents the 75 most frequently occurring keywords in the cleaned corpus in a word cloud, with the relative size of a word being an indication of its frequency. As can be seen by its relative size, “more” is the most frequently occurring word in the corpus used by students in responding to Q2 with over 50,000 instances within the Q2 corpus. In contrast, the second most frequently used word ‘students’ has just over 20,000 instances, and the third ‘lectures’ fewer than 10,000 instances.

[11] Chapter Two contains a more thorough discussion of the steps taken to get raw data into its cleaned state.

Figure 3.5: Relative frequency of the top 75 most frequently used words (Q2)



As has been noted before, word clouds are good at providing an initial indication of the frequency of words used in the corpus but not much beyond this. As such, Figure 3.5 plots the relative frequency of words instead, and reiterates the fact that the word 'more' is by far and away the most popular word in the Q2 corpus.

Table 3.2: Ranking of the top 100 most frequently used words by year

Keyword	2016	2017	2018	2019	2020	Keyword	2016	2017	2018	2019	2020
more	1	1	1	1	1	rather	51	44	40	45	43
students	2	2	2	2	2	information	52	62	54	53	56
work	3	4	4	3	4	learn	53	51	47	48	49
learning	4	5	6	5	5	maybe	54	68	63	61	61
lectures	5	3	3	4	3	groups	55	58	64	51	55
lecturers	6	6	5	7	7	academic	56	67	52	54	57
class	7	7	7	6	6	encourage	57	59	61	56	63
course	8	10	12	11	10	courses	58	50	58	58	62
better	9	9	9	8	8	lecturer	59	65	59	71	66
time	10	12	11	12	11	good	60	54	51	59	60
provide	11	11	10	10	13	first	61	73	73	87	88
feedback	12	8	8	9	9	activities	62	61	53	52	39
student	13	16	16	16	16	emphasis	63	81	79	80	82
make	14	13	13	14	14	sure	64	57	57	41	42
classes	15	15	15	15	15	interaction	65	63	70	66	64
assignments	16	14	14	13	12	available	66	83	67	70	84
year	17	17	23	23	25	different	67	60	65	62	65
help	18	20	22	21	22	continuous_assessment	68	55	56	63	59
give	19	22	20	22	21	exam	69	66	60	60	54
think	20	24	27	26	33	facilities	70	91	72	93	85
tutorials	21	18	18	19	23	understand	71	80	78	84	73
group	22	23	21	17	18	really	72	88	93	92	110
less	23	21	17	18	17	subjects	73	69	99	83	100
college	24	27	35	42	36	interactive	74	48	50	37	38
modules	25	19	19	20	19	difficult	75	71	81	77	77
lecture	26	26	24	24	24	take	76	94	75	85	91
improve	27	25	25	28	28	attendance	77	77	71	73	80
much	28	36	45	43	40	focus	78	89	90	78	79
know	29	29	28	25	26	subject	79	100	101	127	155
study	30	28	26	32	34	well	80	64	83	82	83
practical	31	33	31	27	31	instead	81	75	69	68	67
material	32	37	33	39	35	go	82	111	92	95	90
support	33	35	34	31	32	semester	83	78	82	75	74
library	34	40	39	44	48	many	84	70	91	86	78
exams	35	30	29	33	29	little	85	74	94	137	112
smaller	36	31	30	34	30	certain	86	85	103	108	111
nothing	37	32	32	29	27	skills	87	114	109	103	115
people	38	34	41	35	37	every	88	110	105	124	103
engage	39	41	36	38	46	needs	89	98	117	161	133
feel	40	38	46	55	58	projects	90	101	98	97	86
teaching	41	42	38	36	41	topics	91	87	68	65	93
need	42	43	55	49	50	years	92	79	84	117	127
staff	43	56	66	69	71	find	93	115	118	100	132
experience	44	45	62	64	68	reading	94	76	80	72	72
online	45	39	37	30	20	examples	95	93	77	79	75
hours	46	53	44	46	53	communication	96	117	120	106	87
questions	47	49	48	40	44	often	97	104	114	114	97
engagement	48	47	43	50	51	institution	98	97	124	109	136
module	49	46	49	57	52	placement	99	143	161	138	144
notes	50	52	42	47	45	week	100	86	88	112	108

There are 15,600 unique words used in the clean Q2 corpus, and these words are used 728,000 times in total. However, much like what was observed for Q1, the top 75 words account for over forty percent of the words used in the Q2 corpus.

While we could view the corpus as being a single body of data collected over five years, it can also be viewed as five separate collections of data, each collected in a separate year (2016-2020). These corpora are never going to be identical, as each year brings with it a different cohort of respondents with different experiences, along with some changes in approaches to teaching and learning by their HEIs even if the curriculum remains the same[12].

Nevertheless, it is worth examining the popularity of words in the corpus over time to see if there are any substantial changes or if the popularity of words used by respondents remains stable over time. Table 3.2 presents the ranking of the top 100 words used in each year and has been colour-coded so that high ranked words are green, and lower ranked words are red, with words in-between these poles moving along a continuum from yellow to orange. Changes in the colours over time indicate shifts in ranking, so moving from green towards red means that the word has fallen down the ranking. Conversely, moving from red to green indicates increased usage and a rise in its relative rank.

As can be seen from this table, and somewhat of a contrast to Table 2.2, there is remarkable stability. Words as they go down the ranks appear not to substantively shift across years, thus in the top half of the table, most words remain a similar colour to their initial shade. A similar pattern is evident in the lower half, though towards the bottom we see some red as words decline somewhat in their usage. All of this tends to suggest that changes over time have been minor, and students tend to approach answering Q2 in a similar way as the frequency and ranking of words used in their responses are very similar over time. This is tested further in the next section.

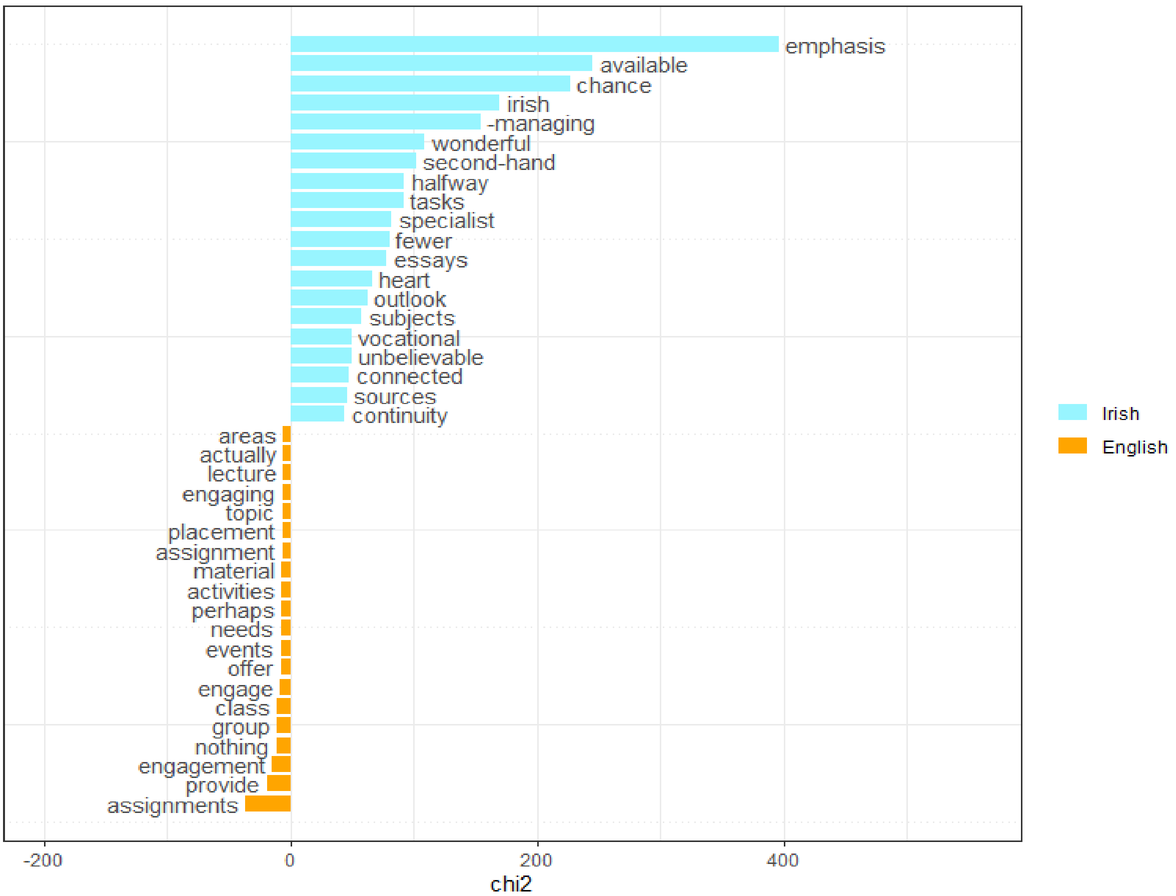
[12] See footnote 2 above for discussion of our implicit assumptions about each year having broadly similar experiences.

3.2 Relative Frequency Analysis (Keyness)

The last table examined the frequency of words used over time; however, it is also possible to compare usage of words across groups. Keyness is the statistical measure to identify significant differences across groups, which uses the relative frequency of words across two parts of a corpus to see if there are differential associations of words used between a target and a reference group.

Figure 3.6 highlights differential usage of words used by students conducting the survey in Irish compared against students completing the survey in English. Thus from the top half of the chart, we can see that students conducting the survey in Irish appear to use the words 'emphasis', 'available', and 'chance' more than their counterparts who completed the survey in English. In contrast, students doing the survey through English marginally use the word 'assignments' more than their Irish language counterparts.

Figure 3.6: Relative frequency analysis (keyness) by language

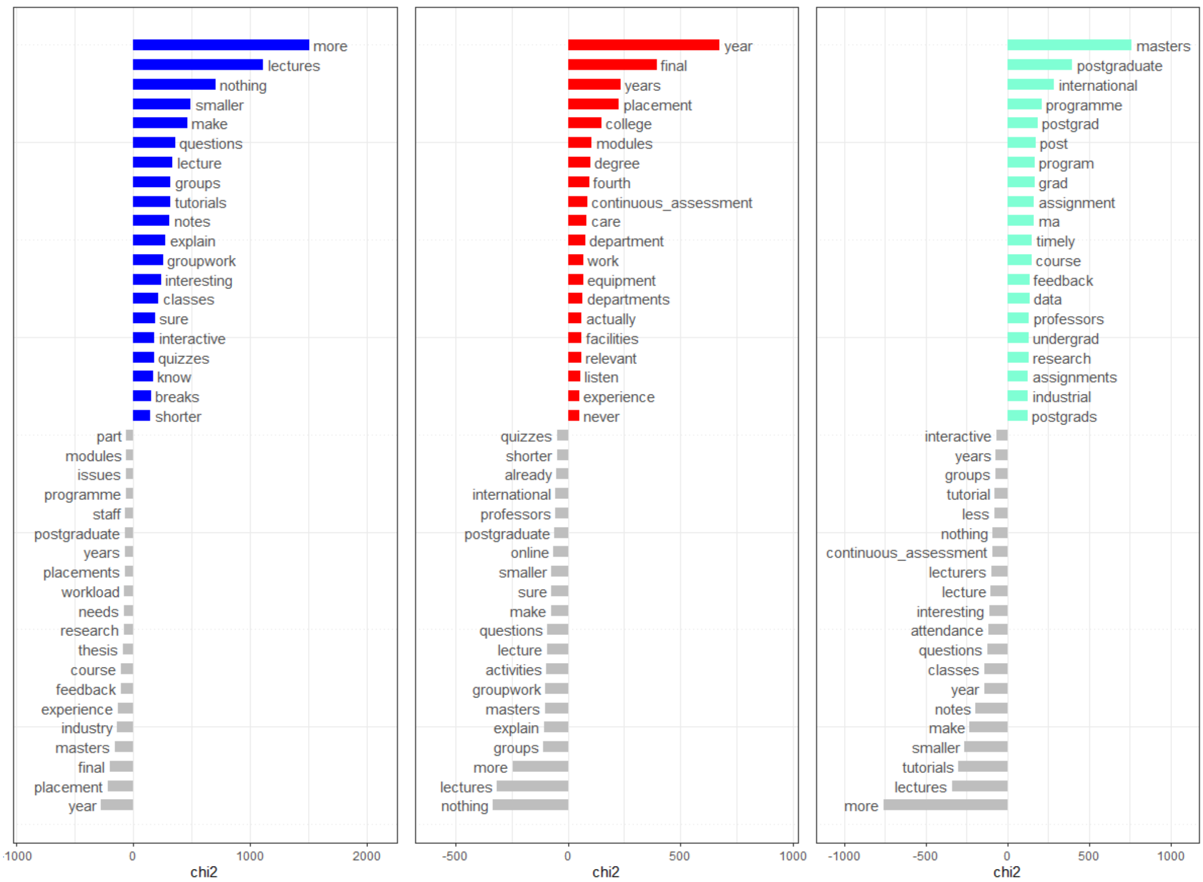


We queried the Irish language differences with our translator and the reason for this difference appears to be that Irish has fewer synonyms than English so the word ‘*béim*’ is frequently used by students and can be translated as emphasis(e), or focus, or direct but for the purposes of this study has been translated as emphasis. A similar pattern is the reason why ‘available’ and ‘chance’ are over-represented in the Irish corpus.

On the other side, the Irish corpus is very small compared against the English corpus, so we are likely to see some differences purely due to the fact that the English corpus is larger and contains a wider range of words than the Irish corpus.

Figure 3.7: Relative frequency analysis (keyness) by student status

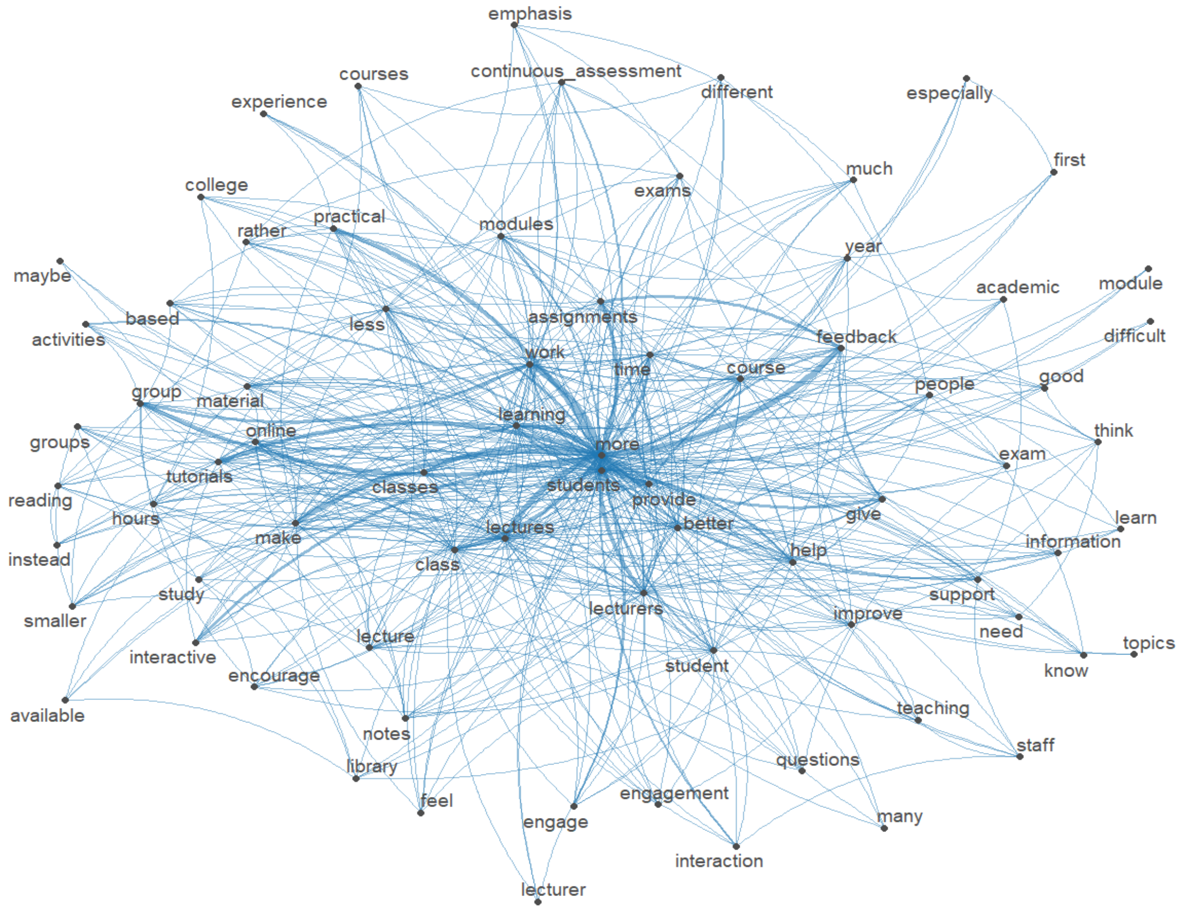
First year undergraduates	Final year undergraduates	Taught Postgraduates



It is also worth examining if the status of a student affects the relative frequency of words used in their responses. Figure 3.7 examines this by comparing each group against a reference group (i.e. the other two groups) to see any relative frequency difference in usage of keywords. The bars in grey are terms frequently used by the reference groups and those in colour are the target group for that chart. The previous chapter found significant differences where final year students mentioned ‘final’ and ‘year’ more than other groups and taught postgraduate students mentioned ‘masters’ more than other groups, and a similar pattern emerges here with students groups tending to mention their own group more than other groups, and in addition mention keywords more commonly associated with their stage in third level education. Thus, first year students mention ‘lectures’ and ‘tutorials’ more than other groups, and final year students mention ‘placements’ and ‘continuous assessment’ more than other groups.

The last chapter discussed the statistical implications of these tests and argued that while these differences are significant they are not substantive, and a case can be made for the same here as these charts tend to show that different groups have subtly different priorities which are expressed through the different frequencies of certain words used, however, these charts do not show huge substantive discrepancies across groups; so each groups’ experiences of student life have more in common with one other than their relative differences.

Figure 3.8: Semantic network of the feature co-occurrence matrix



So far, the analysis has been of individual words, which has shown us the frequency of words used by students at an aggregate level, and some patterns with how these frequencies change when student groups are disaggregated or across time. The next step is then to identify which words are most associated with one another. Within the statistical software this was done by creating a feature co-occurrence matrix which records the number of co-occurrences of tokens.

This feature co-occurrence matrix can then be visualised in a semantic network to illustrate which words are most associated with one another. The width of the bars linking words indicates the strength of the relationship between the words. Figure 3.8 presents a semantic network and as noted already, the frequency that 'more' is found in the corpus means that it is linked with a large number of other keywords and forms the central hub from which all other words branch out from.

This chart broadly illustrates that a large number of students answered Q2 by writing “more” and following this up with a mention of some item that concerned them. All of which, supported our hypothesis that to get a handle on how students answered Q2 it was necessary to examine both the items that were being mentioned but also the degree and direction associated with them (for example, more/less/fewer/better and so on).

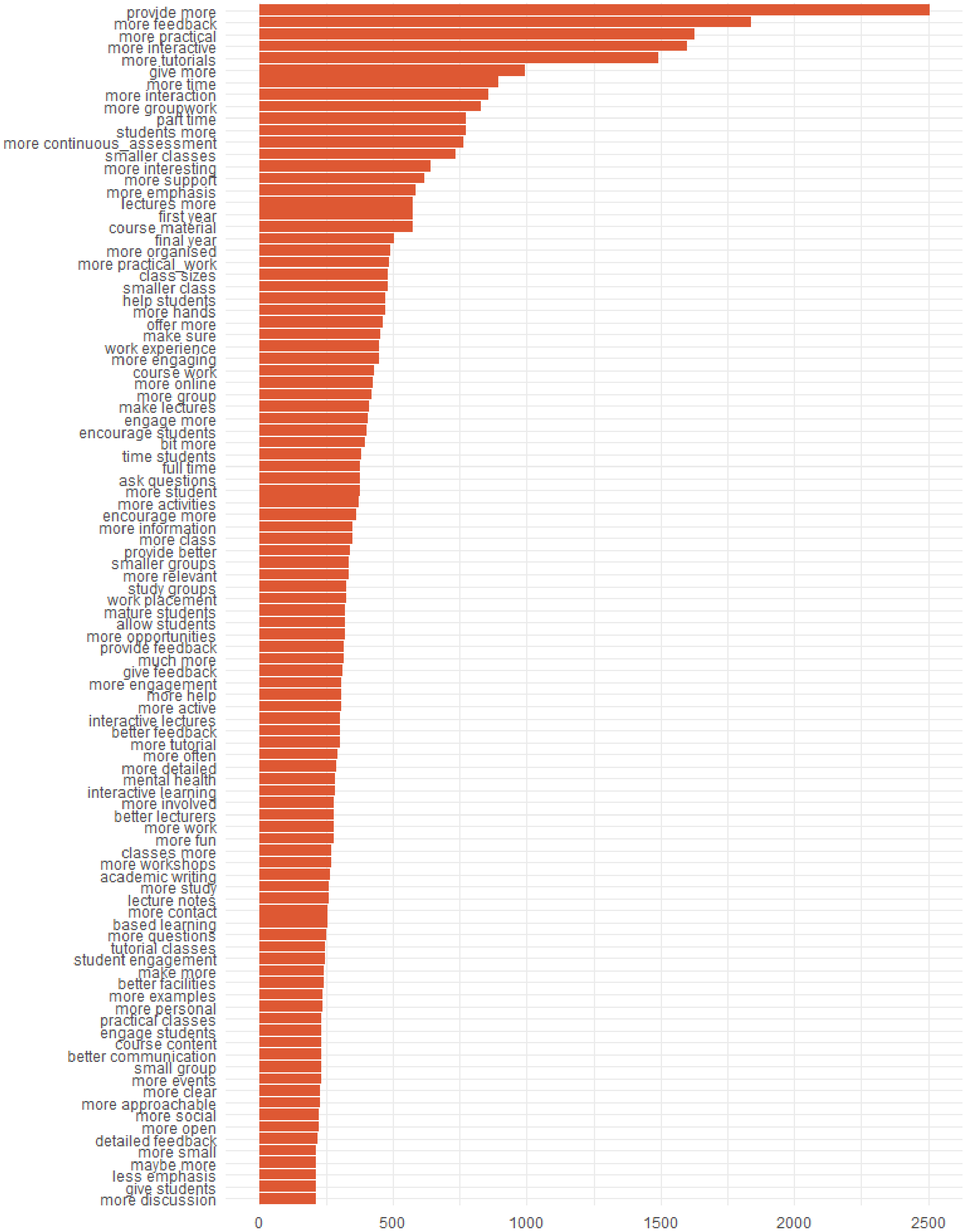
3.3 Bigrams and Trigrams

At this point in the analysis of Q1 we utilised latent semantic scaling to bring in sentiment analysis of words in order to gain an understanding of whether words used by students had positive or negative connotations. This is of less use here where we would expect answers to contain both what students think can be improved by listing an item or items, along with how it can be improved. As a result of this, the positive and negative connotations associated with a modifier is of less utility as the usage of the word encapsulates its definition, more means more, less means less, and so on. Furthermore, the sentiment contained in the comments provided to Q2 are of less use to us in *providing answers* to Q2.

Instead, because students tended to follow a very similar format in answering Q2 through mentioning some form of direction and degree along with an item, it is possible to move beyond individual words in the corpus and tokenise consecutive sequences of words within each comment provided by students, and then examine these sequences as they tend to capture the format with which students provide answers to Q2 very well. These sequences are called n-grams and the ones with greatest use here, are bigrams and trigrams which are sequences of two or three words. So, by seeing how often word X if followed by word Y, and then word Z if interested in the trigram of the sequence, we can then begin to build a model of the relationships between the words, and obtain a picture of the answers students are providing for Q2.

Figure 3.9 presents the most frequently occurring bigrams in the corpus. Some bigrams are completely descriptive such as ‘part time’, ‘mature students’ and ‘work experience’. However, what is striking is how often modifiers are used within the corpus. Throughout this list ‘more’ is often the first or second word in the bigram. Other modifiers include ‘smaller’, ‘better’ ‘less’. These modifiers are sometimes linked with verbs, for example, the most frequently used bigram is ‘provide more’, but more often the modifier is linked to an item, as can be seen in the second most frequently used bigram is ‘more feedback’.

Figure 3.9: Relative frequency of most frequently occurring bigrams



Even when modifiers are linked with verbs, they tend to end with a specific item. Table 3.3 extends the most frequently used bigram 'provide more' to trigram length and shows that 'provide' ironically does not provide us with much in the way of understanding students answers, 'more' is doing the heavy-lifting in the n-gram and is then supplemented by an item be it feedback, support, tutorials and so on.

Table 3.3: Relative frequency of trigrams containing 'provide' as the first word

Rank	Word 1	Word 2	Word 3	Count
1	provide	more	feedback	298
2	provide	more	support	140
3	provide	more	tutorials	112
4	provide	more	opportunities	91
5	provide	more	information	82
6	provide	more	detailed	60
7	provide	more	practical	53
8	provide	more	help	47
9	provide	more	learning	36
10	provide	more	resources	33
11	provide	more	online	32
12	provide	more	time	31
13	provide	more	workshops	30
14	provide	more	study	29
15	provide	more	facilities	28

As such, this simple construction of modifier plus item appears to be a simple and precise way of capturing students' answers to Q2. To demonstrate this further, Figures 3.10 to 3.15 present the frequency of bigrams with the common modifiers 'more', 'less', 'better', 'fewer', 'improve' and 'reduce' as the first word in the pair. There are a few points to note in these charts. Firstly, due to how Q2 is worded, there appears to be a natural inclination to providing a positive response, as a result, positive modifiers such as 'more', 'better', and 'improve' are more commonly used than the negative modifiers such as 'less', 'fewer', and 'reduce' as can be seen in the number of bigrams along the x-axis. Secondly, note how often the bigrams concisely summarise students' main points in their responses. Though it is also worth noting that some bigrams remain vague because the bigrams do not end in an item, for example, 'more practical' in Figure 3.10 or 'less emphasis' in Figure 3.11. These are discussed further in the next section as bigrams are combined with Markov chains.

Figure 3.10: Relative frequency of bigrams containing ‘more’ as first word

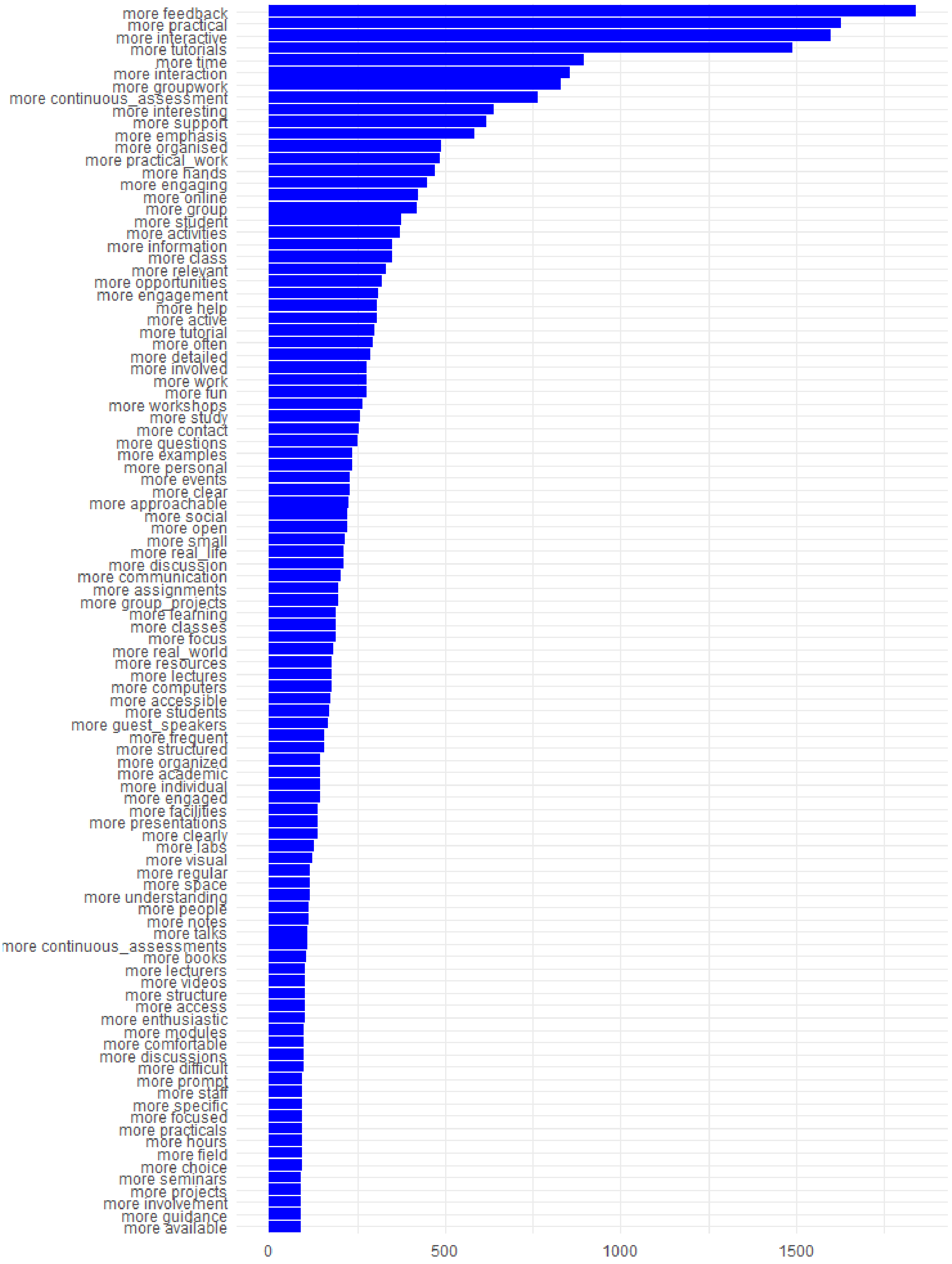


Figure 3.11: Relative frequency of bigrams containing 'less' as first word

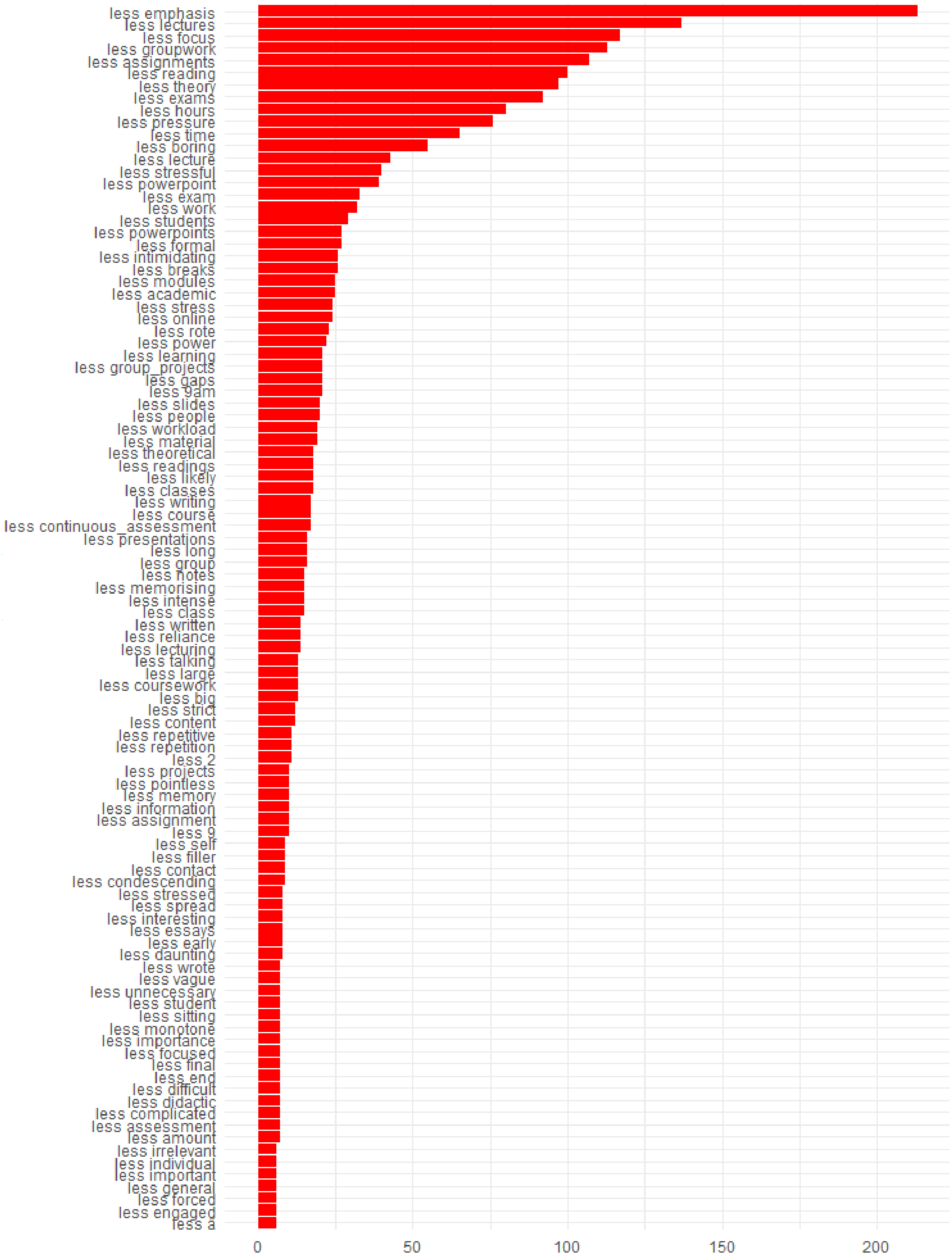


Figure 3.12: Relative frequency of bigrams containing ‘better’ as first word

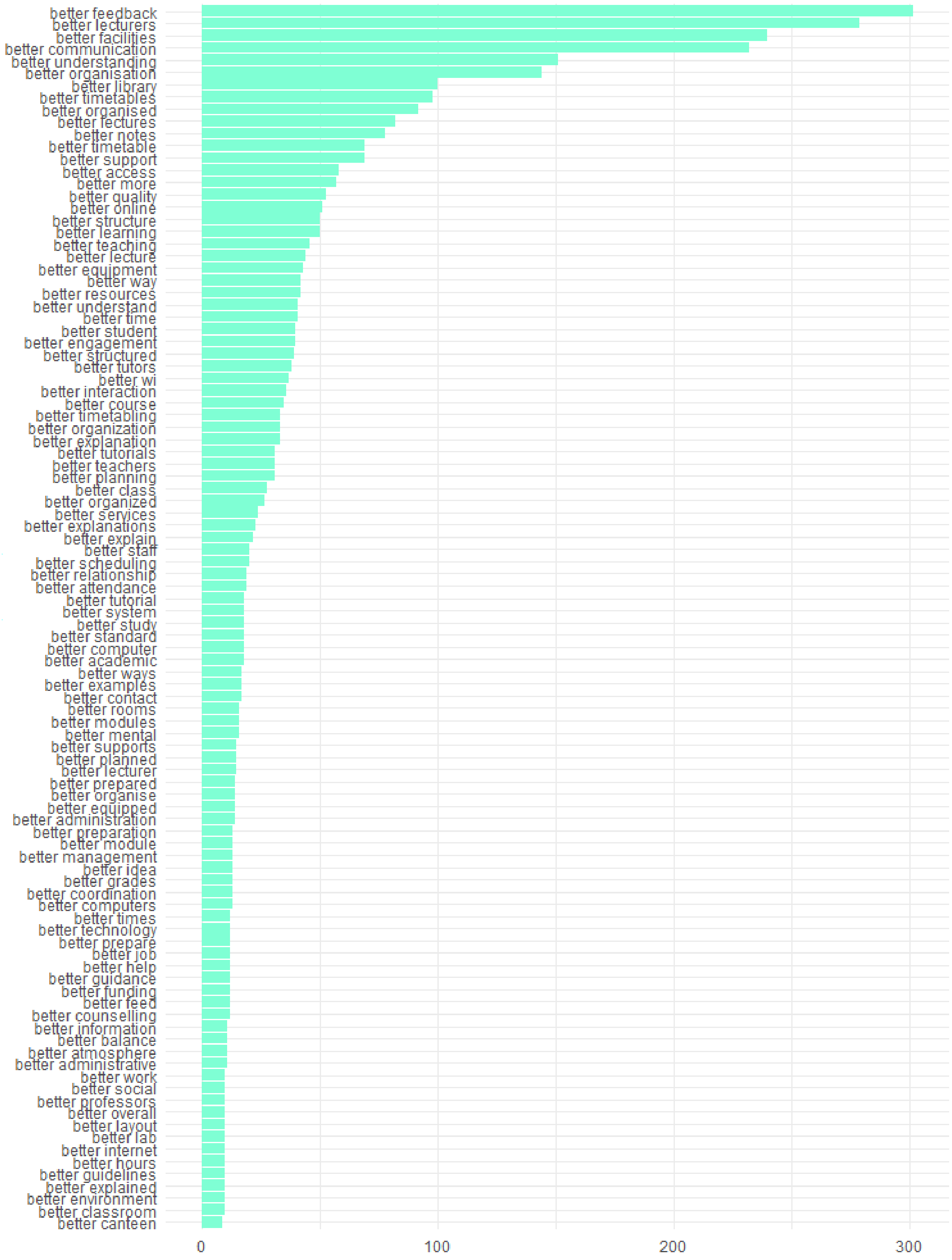


Figure 3.13: Relative frequency of bigrams containing ‘fewer’ as first word

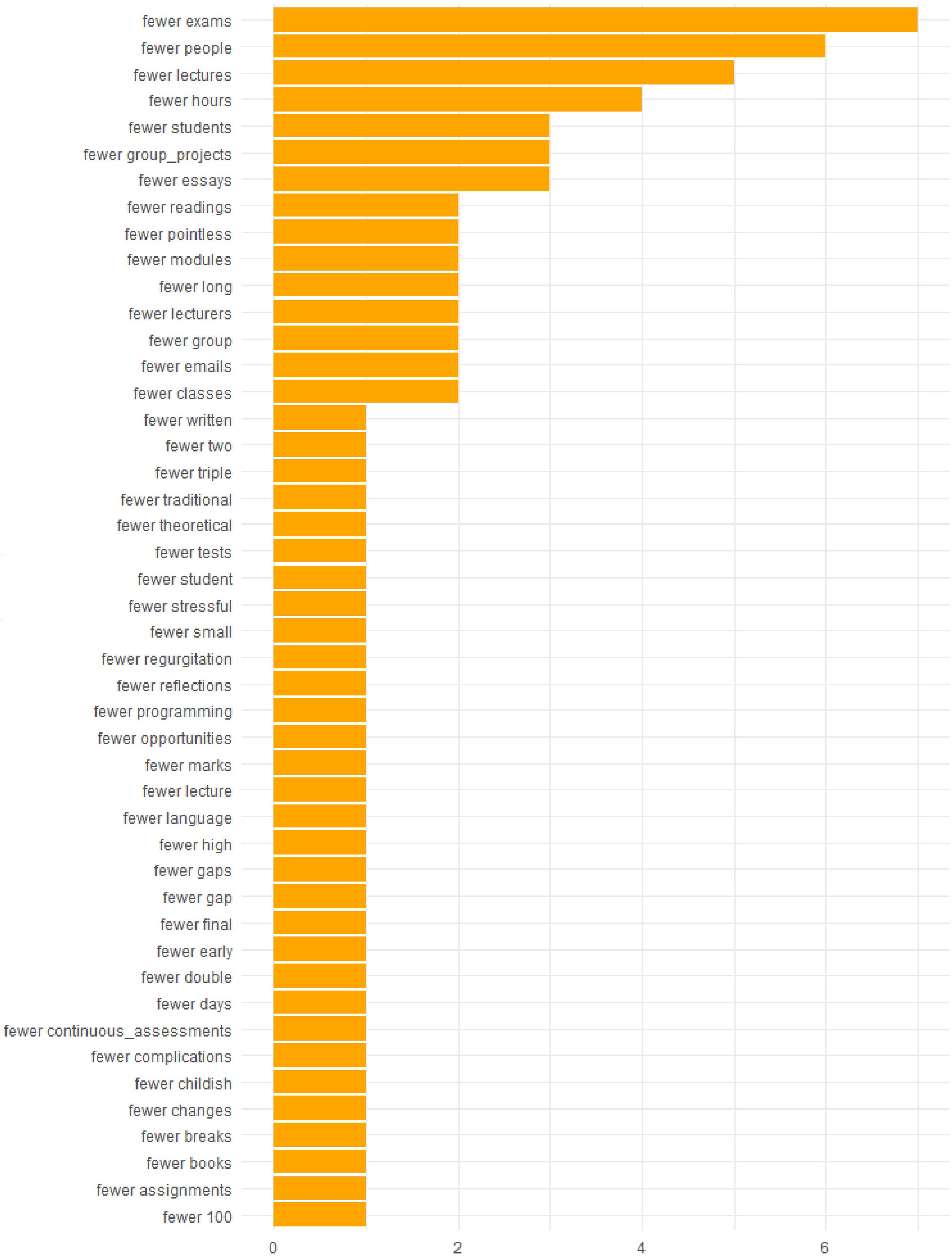


Figure 3.14: Relative frequency of bigrams containing ‘improve’ as first word

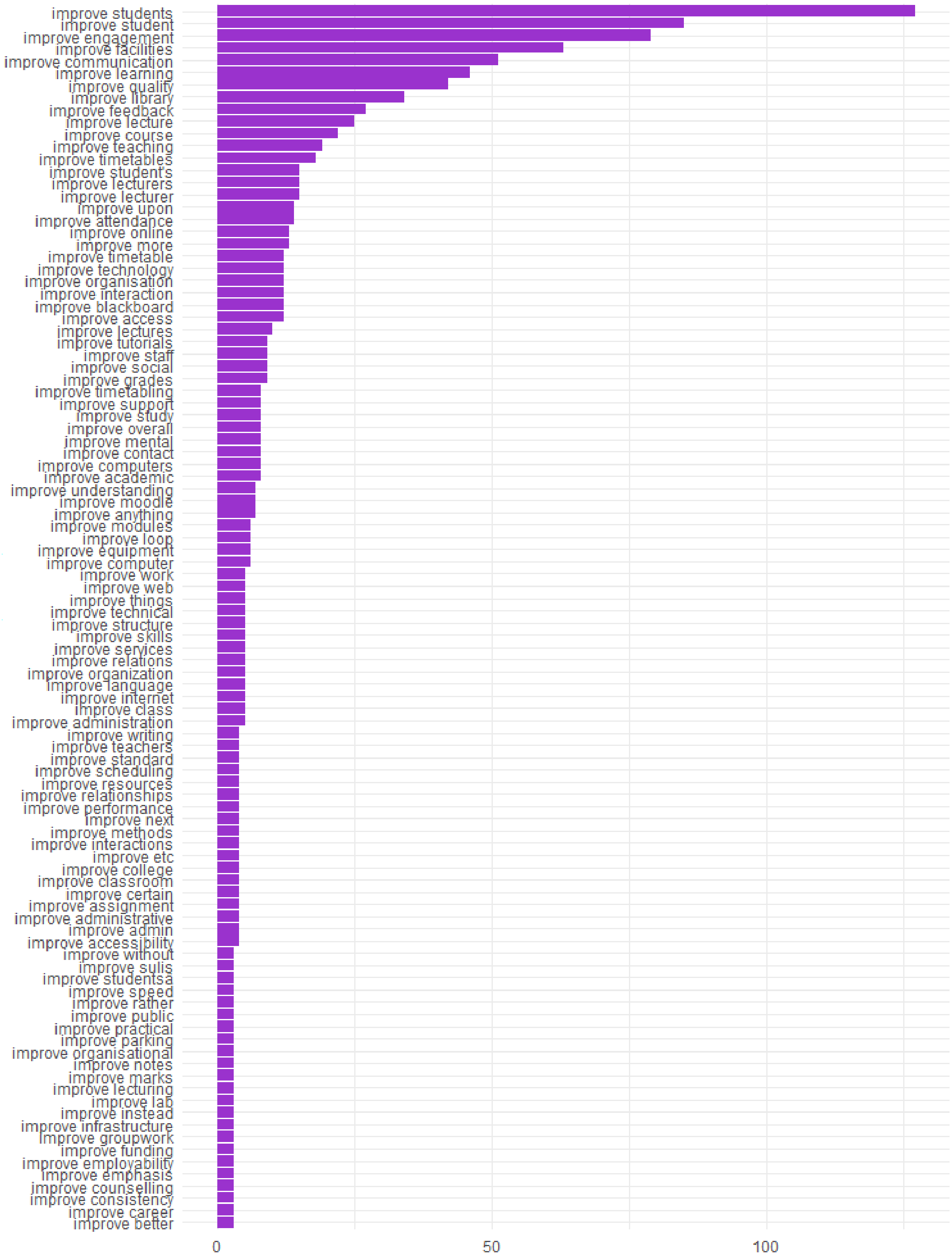
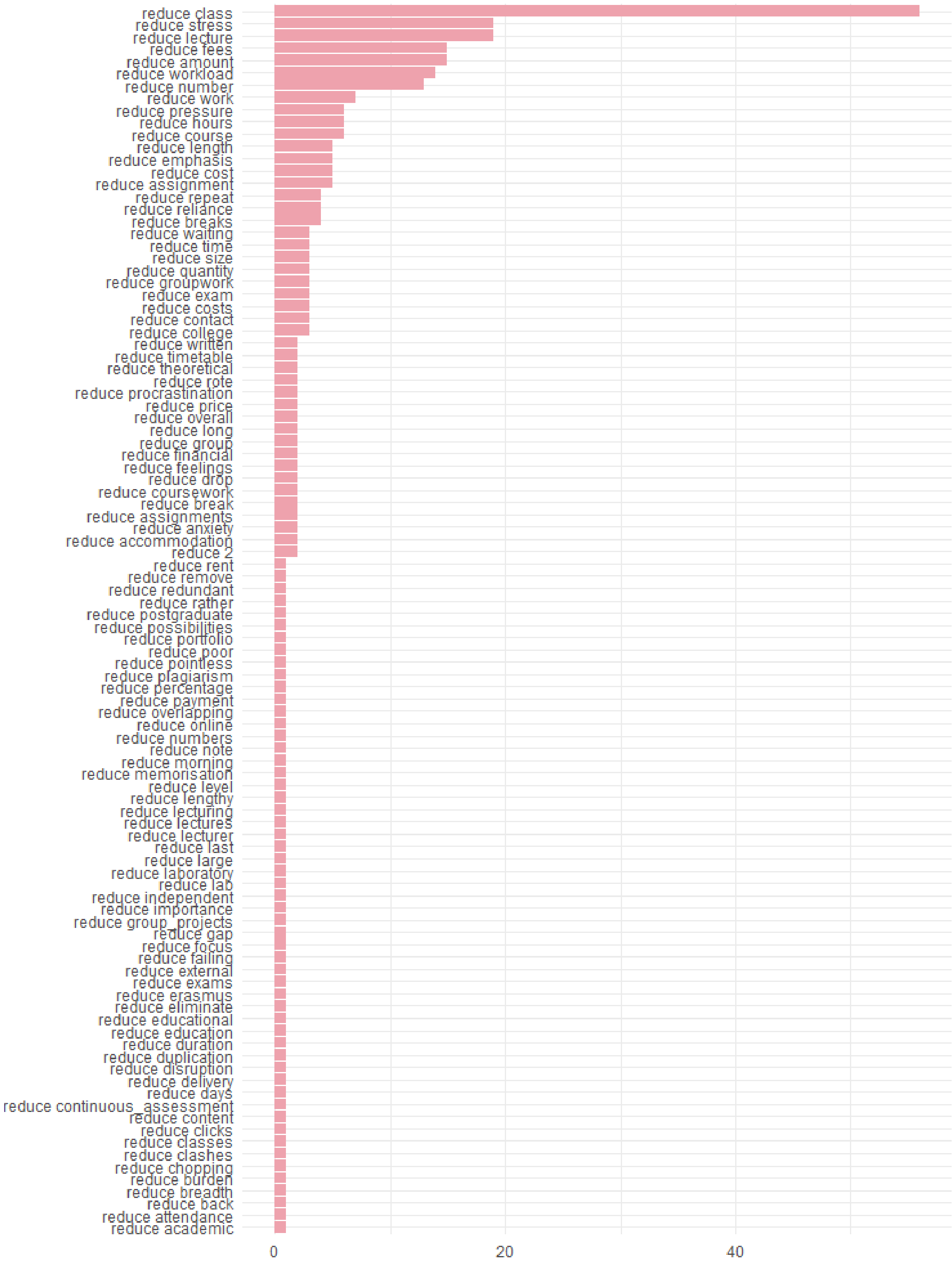


Figure 3.15: Relative frequency of bigrams containing ‘reduce’ as first word



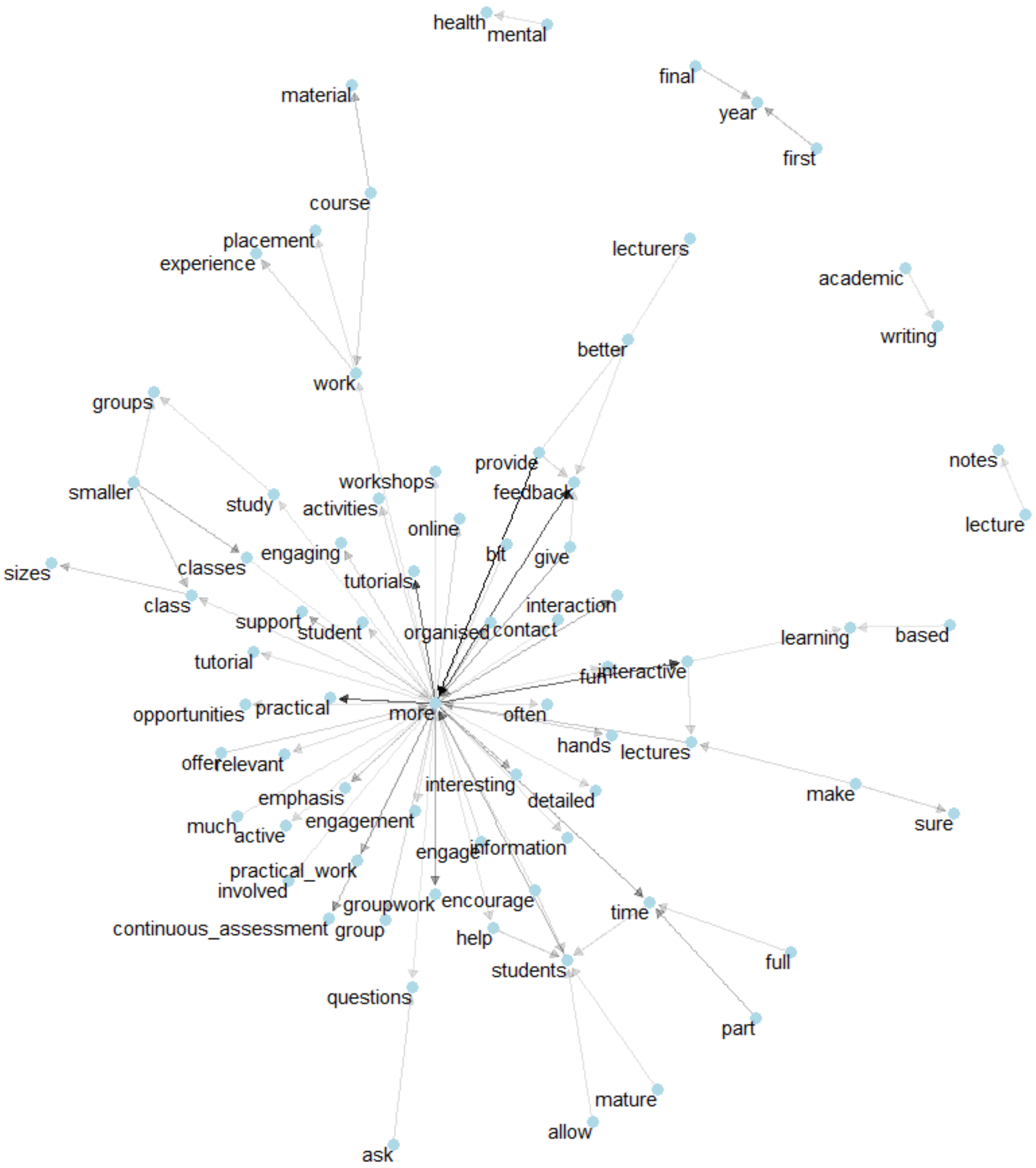
3.4 Bigram Networks

The analysis conducted so far has shown that responses to Q2 tend to follow a formula of modifier plus item and has shown how effective bigrams are in outlining the responses of students. Presenting the frequency of certain modifiers allows us to see how students have responded but requires us to be selective; what word are we interested in, and do we want to use the word as word 1 or word 2 in a bigram? However, doing this is only ever going to provide us with a snapshot of responses that use the terms we are interested in. This is not in itself a poor choice of direction in analysis, but it is also not going to provide any answers about how students answered Q2 as a whole. The next step is to see how the bigrams relate to each other rather than just viewing them in isolation from one another.

The charts presented in the section show bigrams where there are at least 250 cases and visualise the relationships among words simultaneously, rather than just a selected word, and words frequently associated with each. As such, the charts are a visualisation of a Markov chain, which is a common model in text processing, where the choice of a word only depends on its previous word. The arrows linking words show the direction of association. For example, the words ‘full’ and ‘part’ are commonly associated with ‘time’, and ‘time’ is associated with ‘students’. Examining the links between words through following the arrows builds commonly used phrases within the corpus.

Figure 3.16 presents the Markov chain bigram network for all students and as we would expect at this point with everything that has been covered so far, ‘more’ is the central node from which most words emanate, and repeat bigrams seen already in Figures 3.9 and 3.10 such as ‘more feedback’, ‘more continuous assessment’, and ‘more tutorials’. Beyond this, there are a few words on the right-hand side of the chart which stand apart from the hub around ‘more’, meaning they are most commonly associated with each other than words in the hub. These include ‘academic writing’, ‘mental health’, ‘lecture notes’, and ‘first/final year’.

Figure 3.16: Markov chain bigram network (all students)



Throughout this study, we have repeatedly checked if student status has an effect on answers words used, and answers provided to each of the questions at hand. As such, it is worth evaluating here if student status has an influence on the bigrams used, as we would expect that different student groups to have different priorities based on where they are within their time in third level education.

Figures 3.17 to 3.19 provides the bigram networks for first year, final year undergraduates, and taught postgraduates separately. These networks show remarkable similarities, for example, ‘more’ remains the central node for each group, and each touch upon topics common to all: continuous assessment, feedback, and tutorials, among others. There are also subtle differences demonstrating that different student groups have different priorities. For example, final year students mention mental health, and work experience/placement while these are absent from the other groups, and taught postgraduates mention mature and international students, which are not covered by undergraduates.

Figure 3.17: Markov chain bigram network (first year undergraduate students)

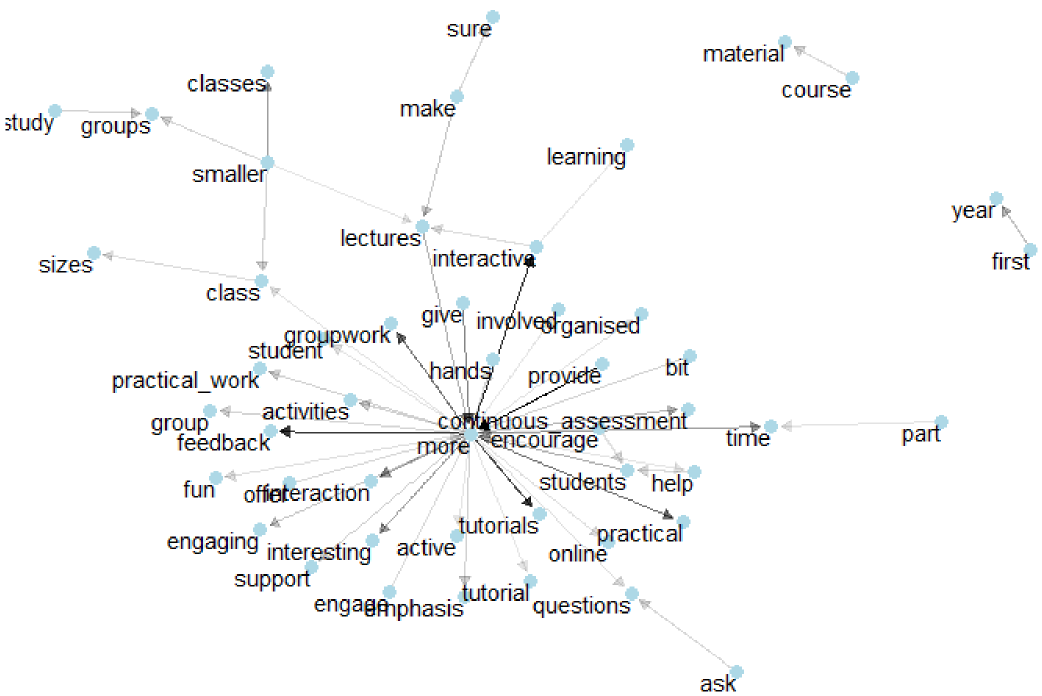


Figure 3.18: Markov chain bigram network (final year undergraduate students)

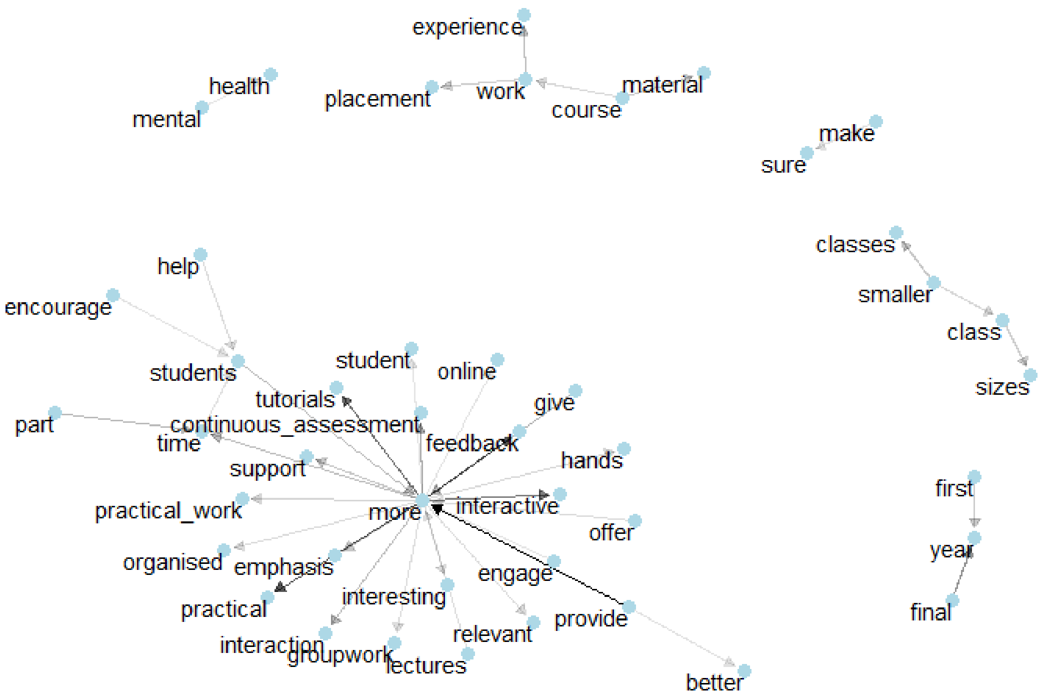
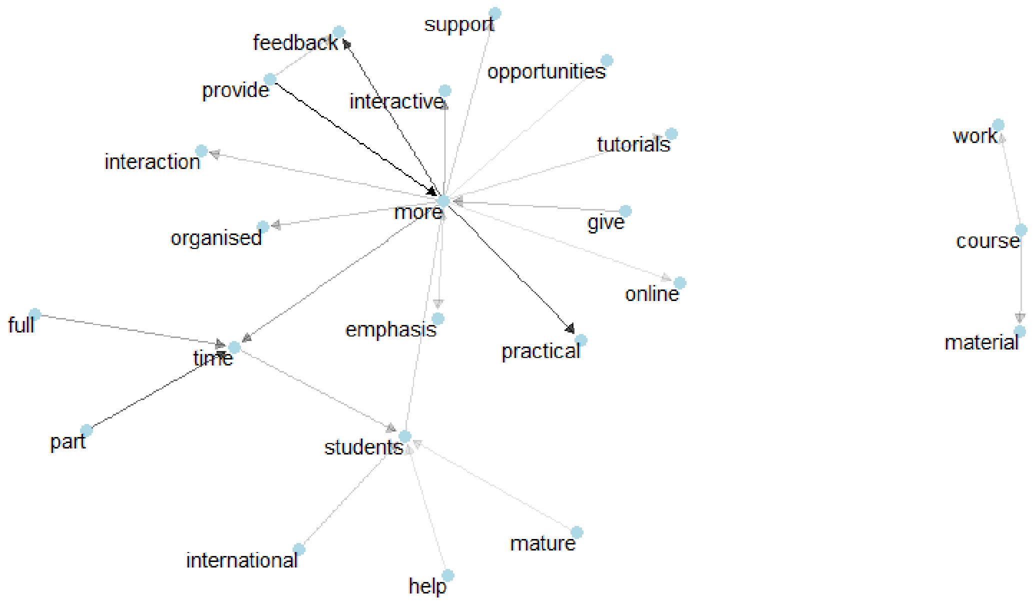


Figure 3.19: Markov chain bigram network (taught postgraduate students)



Taken together these bigram networks present an overview of topics discussed by students in a broader context than the frequency of individual bigrams was able to provide. Though in these networks all bigrams over a certain level (greater than 250 cases) are present in the networks but beyond this, they all are treated equally. Thus, we cannot see how often certain bigrams or topics appear in relation to each other. The final step in this chapter is to examine how often our modifier words of interest are used against topics that repeatedly recur throughout the corpus. As such gaining a more complete picture of what students would actually like their institution to change to increase and improve student engagement.

3.5 Coding Student Responses and Identifying Modifiers

As noted above the final step in the analysis for Q2 is the examine the corpus of student responses and to examine the modifiers used in relation to a prescribed set of items that we already know that students discuss a lot, which can be taken as an indicator of their importance to students in general. As such, this section of the chapter evaluates whole sentences, the keywords of interest contained in them, and the modifiers used in relation to these keywords.

In contrast to Q1 where the question posed gave respondents the freedom to interpret the question as they wished, which meant that there was a wide variety of answers, the structure of Q2 "what could your institution do to improve student engagement?" means that most students has approached answering the question in a similar manner by identifying an item they think could be improved and providing a modifier such as more, less, better, fewer to describe the change that needs to be made to improve student engagement.

This simplifies the analysis somewhat as it makes it easier to grasp what students are telling us regardless of the volume of information provided. While conducting the preliminary review of students' comments we noted themes, and concurrent keywords associated with these themes, which were repeated throughout the corpus. The coding framework provided in Table 2.3 in the previous chapter provides an overview of categories mentioned through the corpus. For consistency, the analysis below was run on all of the keywords, but some categories, especially keywords relating to personal and experiential themes were articulated less often in the Q2 corpus. As a result, the keywords discussed below only provides the results for keywords where there were over 2,000 mentions across all student groups. Table 3.5 presents these topics and the number of occurrences for each across all students and by student status.

Within the statistical software (R), we designed an algorithm that searches for each of these keywords and iterations thereof, and records where they occur. Of the over 91,000 comments recorded by students there were almost 57,000 cases where at least one of the keywords were mentioned. The analysis presented in the remainder of this section is based on this subset of students[13].

A second algorithm was designed to search for modifiers within sentences and note where they occurred and the direction they took. Before running this algorithm to ensure we had a comprehensive list, we went through the 10,000 most frequently occurring words and compiled a list of the modifiers used in the corpus. These are listed in Table 3.4 with the direction they tend to take.

Table 3.4: List of modifiers and direction

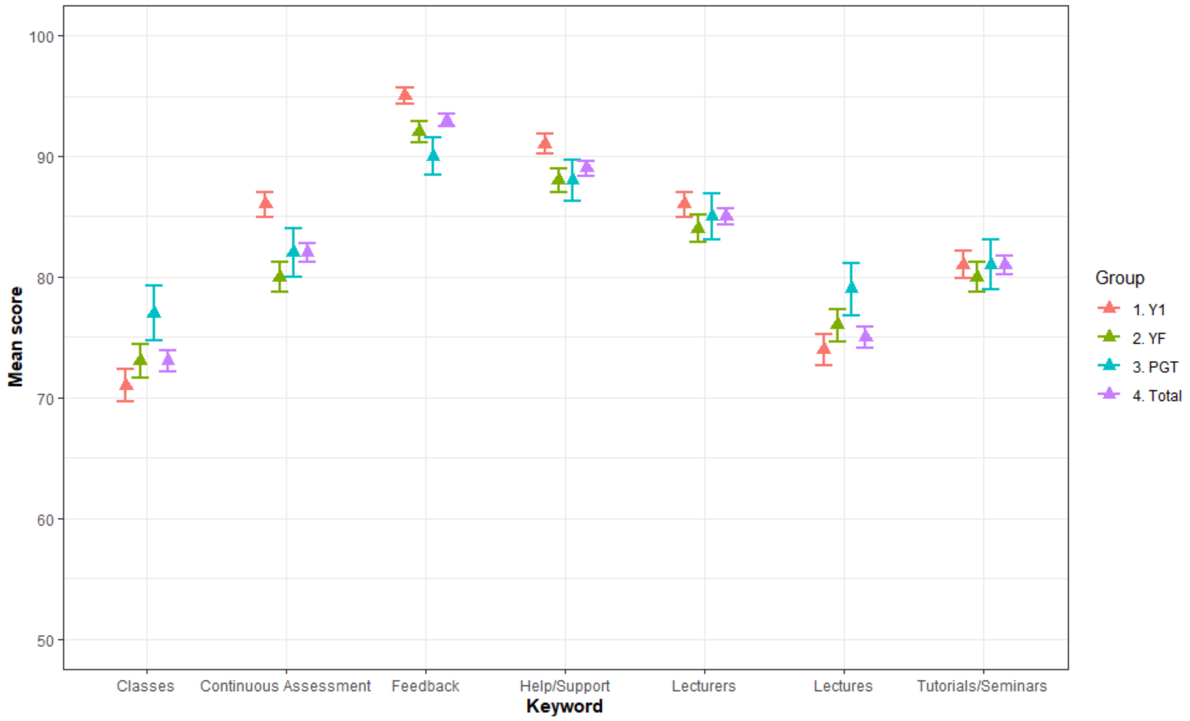
Direction	Modifiers
Positive	more; better; improve; increase; include; greater; promote; higher; effective; flexible; raise; efficient; start.
Negative	less; fewer; reduce; decrease; stop; exclude; small/er; shorter; ineffective; remove; inflexible; lower; inefficient.

Through running these two algorithms we obtained a dataset that provided the number of instances each keyword was mentioned in the corpus and within this, the count of the modifiers used in the sentences that contained each keyword. An overall indication of how students evaluate each keyword can be taken from the sum of positive modifiers over the

[13] Approximately sixty percent of comments contained at least one of the keywords provided in Table 3. Much like the analysis in Chapter 2, this is a very good return for such a constrained set of text. Theoretically, it would be possible to code the entire corpus so that all comments fall into at least one category, but it is not particularly practical and beyond a certain point is of limited utility. However, the scope to extend the coding beyond that already outlined to other areas of interest is available to researchers through the R code provided separately.

total number of modifiers used multiplied by one hundred. This provides a percent score ranging from zero where all modifiers used about a keyword are negative to one hundred where all modifiers used about a keyword are positive. Table 3.5 provides the mean percent score for each keyword for all students and across group. These scores are also presented visually in Figure 3.20. From the chart it is quickly evident that students, regardless of the keyword mentioned, tend to use more positive modifiers than negative ones as all scores are above 50. Furthermore, this chart does not account for the frequency of mentions, as such 'Feedback' and 'Continuous Assessment' share similar scores despite feedback being mentioned twice as much as continuous assessment by students.

Figure 3.20: Average scores of keywords (with 95% confidence intervals) across student status



One assumption about the answers provided to Q2 is that students mention the items that they think could improve student engagement, as such, to gain an overall picture of how students **as a whole** answered this question it is necessary to account for the relative frequency of some keywords over others.

This has been done in the subsequent analysis through weighting of the keywords based on the formula below:

$$\text{Weight} = \frac{\text{Number of Mentions of Keyword (by group)}}{\text{Maximum Number of Mentions of All Keywords (by group)}}$$

To demonstrate using an example from Table 3.5, first year students for the keyword 'classes' have a mean of 71 percent, and this is mentioned by students 4,248 times. The maximum number of mentions for this group (first year undergraduates) is 'lectures' with 4,436.

$$\frac{4{,}248}{4{,}436} = 0.96 \rightarrow 0.96 * 71 = 68$$

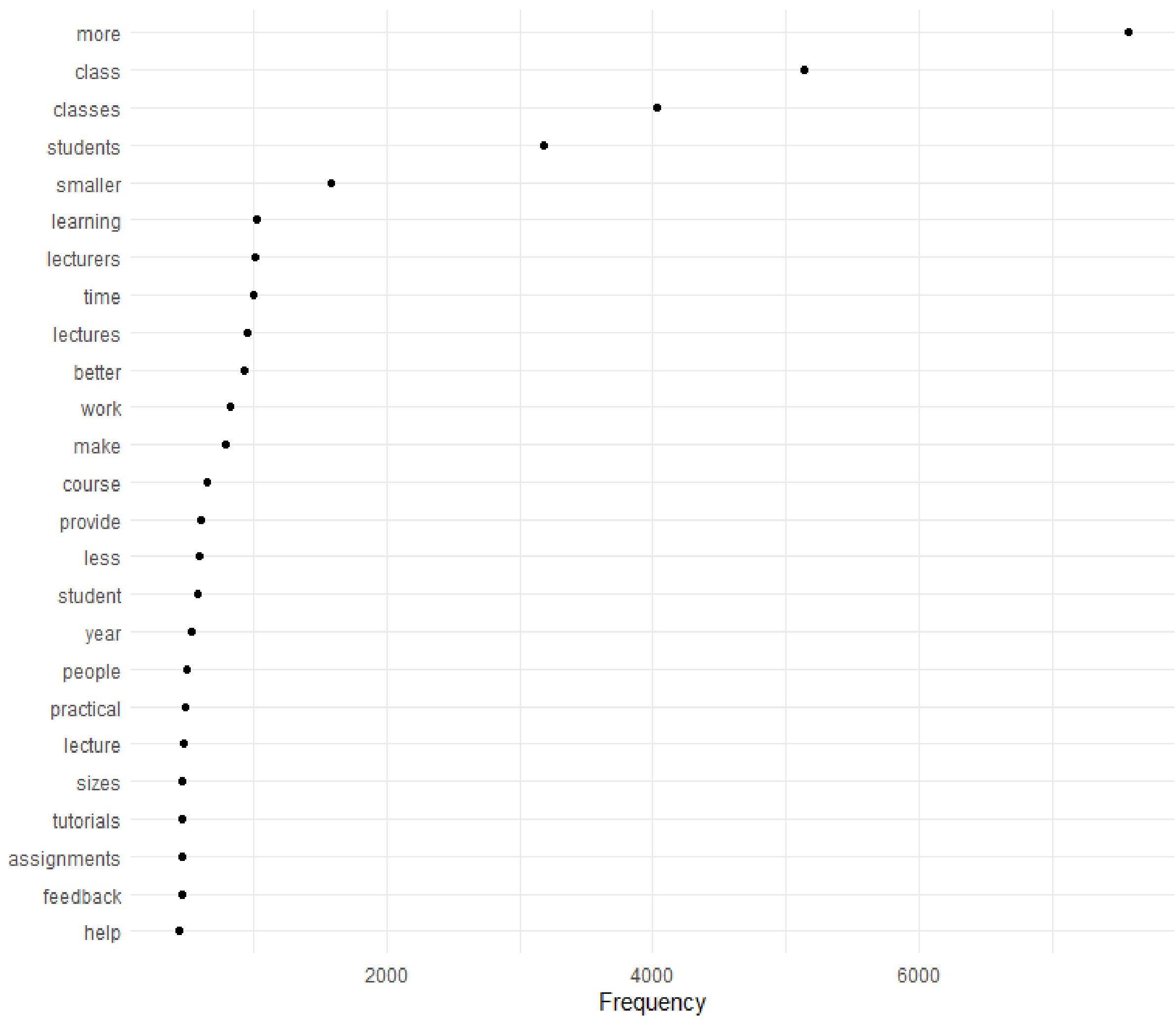
Table 3.5: Keywords, number of mentions, and mean positive-negative scores

Keyword	Group	Mean Percent	Number of mentions	Weighted Mean Percent
Classes	1. Y1	71	4,248	68
Classes	2. YF	73	3,954	73
Classes	3. PGT	77	1,367	77
Classes	4. Total	73	9,569	73
Continuous Assessment	1. Y1	86	739	15
Continuous Assessment	2. YF	80	1,124	22
Continuous Assessment	3. PGT	82	279	16
Continuous Assessment	4. Total	82	2,142	18
Feedback	1. Y1	95	1,591	34
Feedback	2. YF	92	1,988	46
Feedback	3. PGT	90	1,103	73
Feedback	4. Total	93	4,682	46
Help/Support	1. Y1	91	2,133	44
Help/Support	2. YF	88	2,280	51
Help/Support	3. PGT	88	1,012	65
Help/Support	4. Total	89	5,425	51
Lecturers	1. Y1	86	2,438	47
Lecturers	2. YF	84	2,710	58
Lecturers	3. PGT	85	853	53
Lecturers	4. Total	85	6,001	54
Lectures	1. Y1	74	4,436	74
Lectures	2. YF	76	2,708	52
Lectures	3. PGT	79	930	54
Lectures	4. Total	75	8,074	63
Tutorials/Seminars	1. Y1	81	2,254	41
Tutorials/Seminars	2. YF	80	1,811	37
Tutorials/Seminars	3. PGT	81	417	25
Tutorials/Seminars	4. Total	81	4,482	38

4,248 over 4,436 gives a weight of 0.96 which when multiplied by the mean of 71 percent gives us a weighted score of 68 percent. The weighting in effect down-weights less frequently mentioned keywords, but because the weighting is based at the group level it allows for any intra-group variation to become more visible.

Finally, the analysis below was conducted on individual subsets of the dataset created by the algorithms. This reduces the likelihood of other keywords influencing the results of individual keywords. So, for example, the analysis on the keyword 'classes' uses only the portion of the corpus that specifically mentioned 'classes', which reduces the chance that the modifiers in the text refer to another keyword. This reduces the chance of measurement error but does not eliminate it and forms the reasoning behind the need for 95% confidence intervals around each point as the 'true' value of the parameter should be contained within the band.

Figure 3.21: Most frequently used words in the 'classes' subset



With regard to the subset of the corpus containing references to classes, Figure 3.21 shows the most frequently used words and as we would expect at this point, 'more' is the most frequent, 'class' the second, and 'classes' the third. Figure 3.22 presents the weighted mean scores for the proportions and the overall score for all students is around 75 percent with the other groups varying slightly around this point.

'Classes' from our analysis is the only keyword where different directions of modifiers could be used for there to be an overall improvement in student engagement. Most students appear to want more classes. However, as one can see from Figure 3.21, smaller is the fourth most used word which is a negative modifier and is generally used in the sense of students wanting to have 'smaller classes'.

Figure 3.22: Average weighted scores of mentions of 'classes' (with 95% confidence intervals) across student status

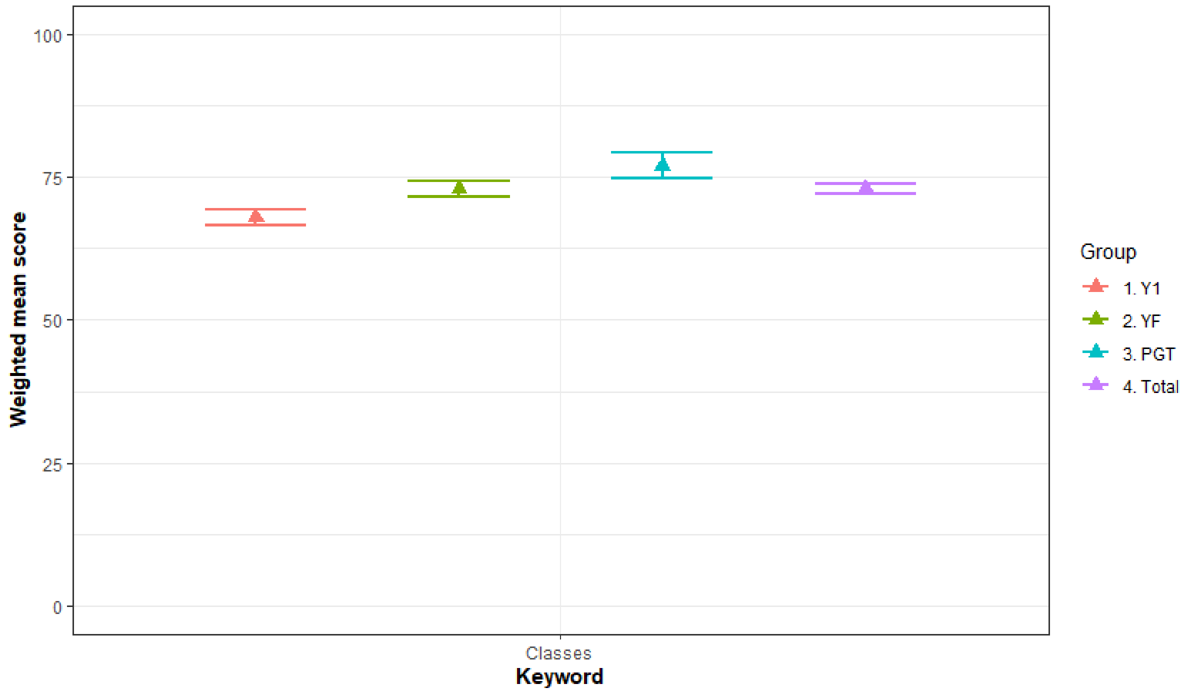
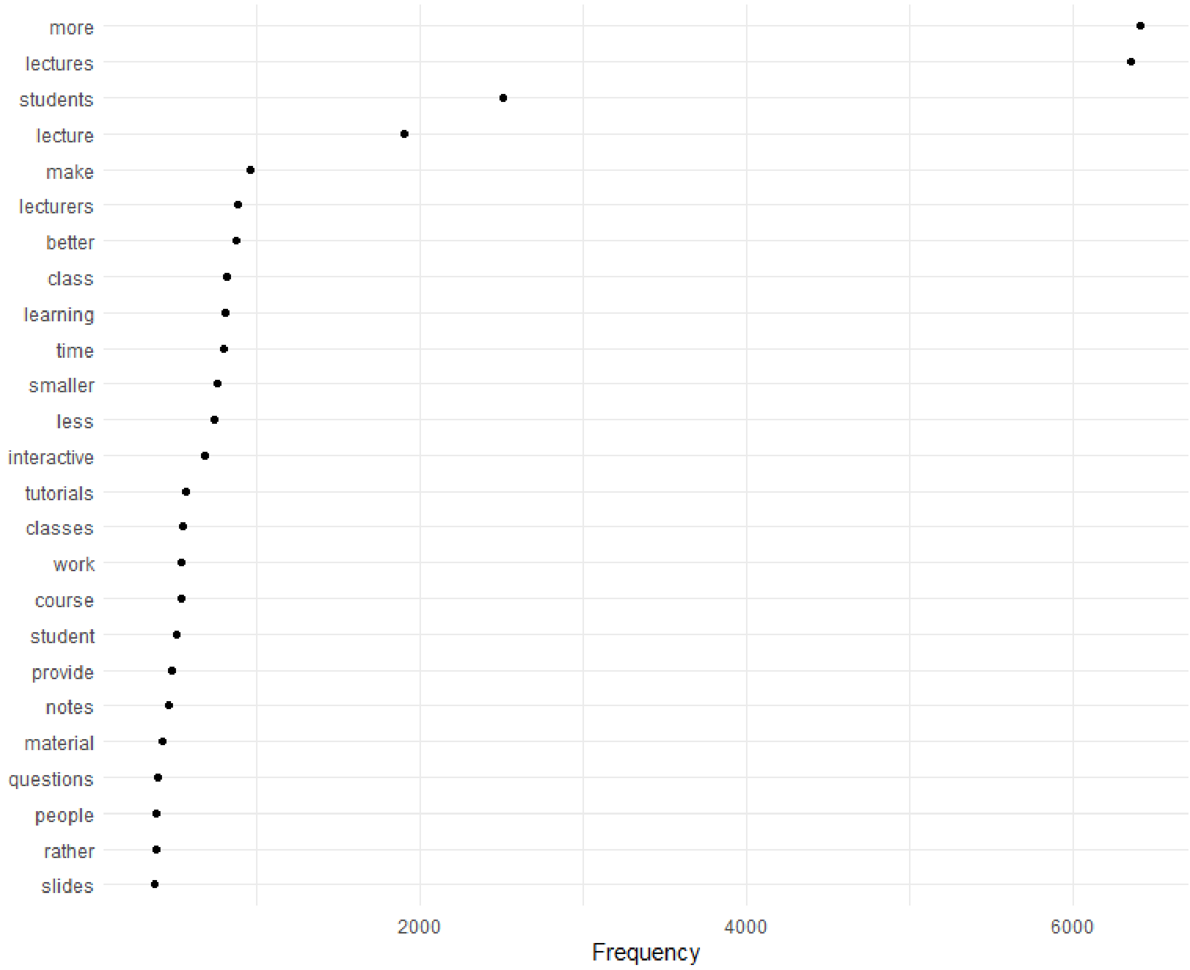


Figure 3.23: Most frequently used words in the 'lectures' subset



With regard to the subset of the corpus containing references to lectures, Figure 3.23 shows the most frequently used words and again, 'more' is the most frequently used word, closely followed by 'lectures'. Figure 3.24 presents the weighted mean scores and the overall score for all students is around 62.5 percent. There are statistically significant differences here as first year undergraduates tend to want 'more' and 'better' lectures than final year and taught postgraduates. The difference is in part because of the weighting, as a greater number of first years mentioned lectures than the other groups, but this also can be taken as an indication that this is a higher priority for first years than final years or taught postgraduates.

Figure 3.24: Average weighted scores of mentions of 'lectures' (with 95% confidence intervals) across student status

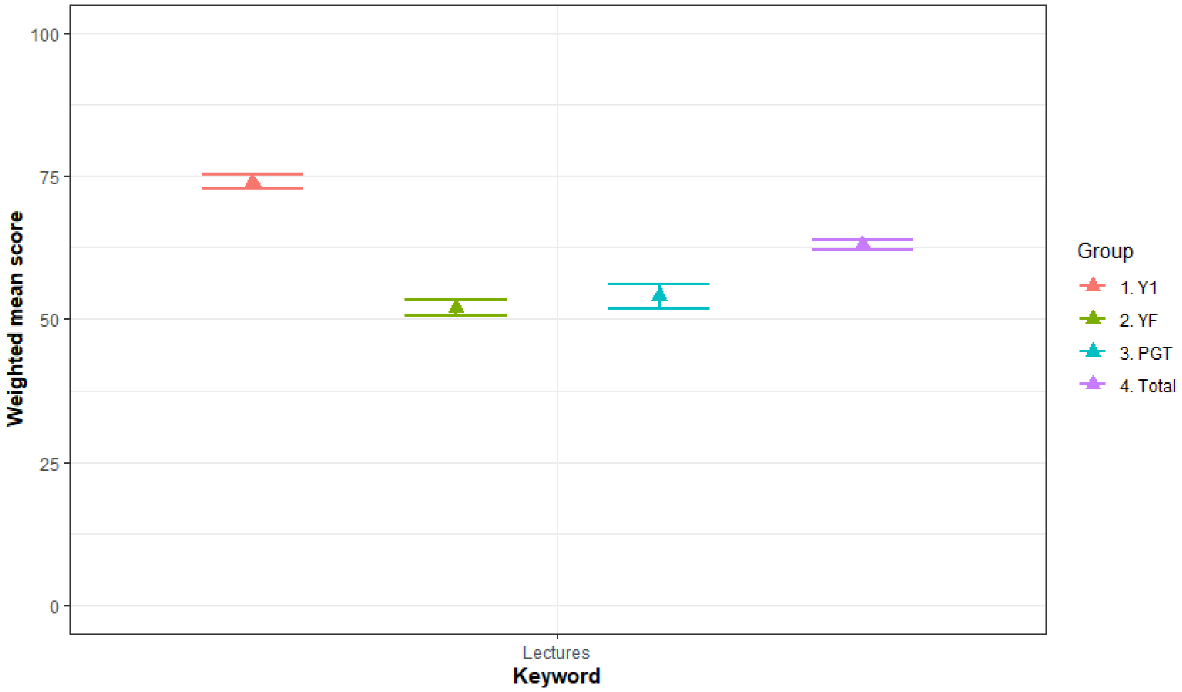


Figure 3.25 shows that in contrast to almost every other keyword bar 'feedback' discussed further below, 'lecturers' is the most frequently used word within the 'lecturers' subset rather than 'more' which is the second. Figure 3.26 presents the weighted mean scores and the overall score for all students is around 50 percent. There are statistically significant differences here as final year undergraduates tend to want 'more' and 'better' lecturers than first year undergraduates.

Figure 3.25: Most frequently used words in the 'lecturers' subset

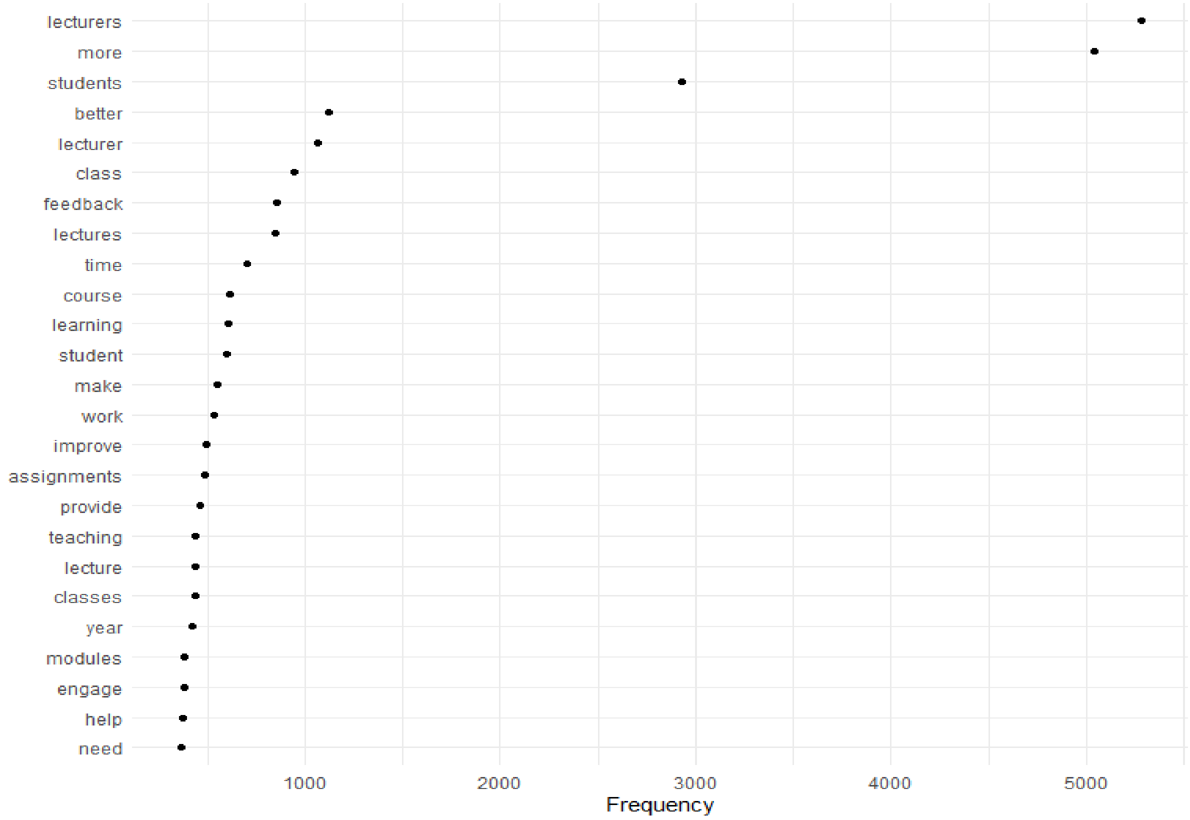


Figure 3.26: Average weighted scores of mentions of 'lecturers' (with 95% confidence intervals) across student status

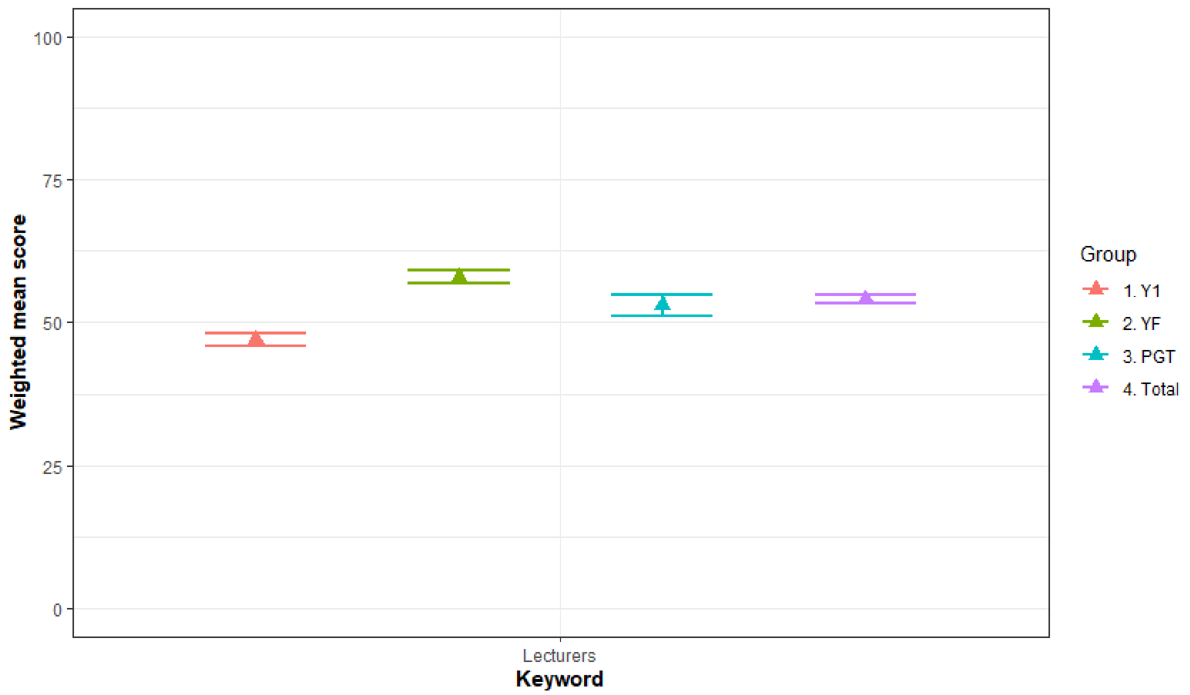
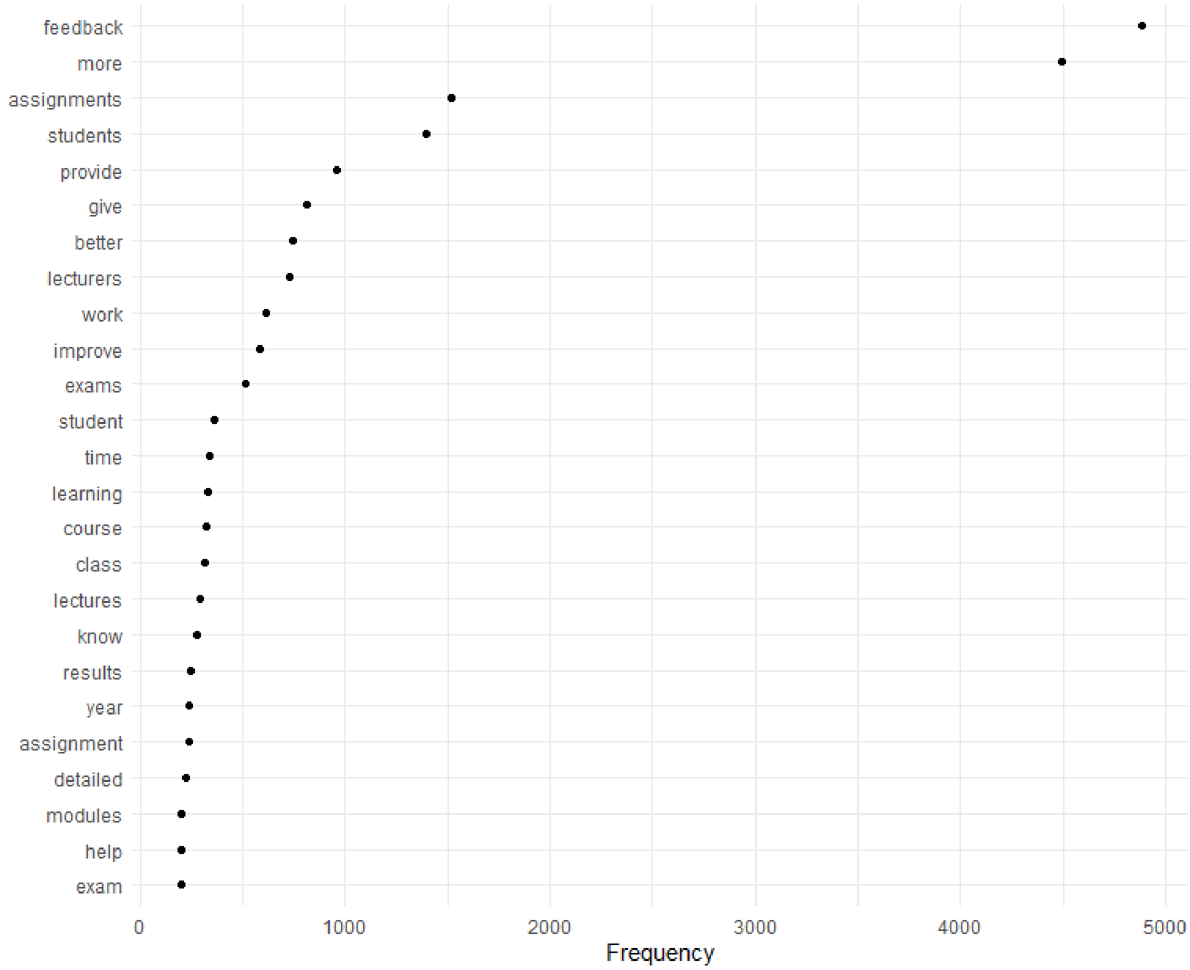


Figure 3.27: Most frequently used words in the ‘feedback’ subset



Feedback like ‘lecturers’ is the only other keyword that appears in its own subset more than ‘more’, as shown in Figure 3.27. In second place is ‘more’ and third is ‘assignments’. Figure 3.28 presents the weighted mean scores for the proportions and the overall score for all students is below 50. However, this chart shows significant differences across groups as taught postgraduates mention wanting to receive more and better feedback on their assignments than the other student groups and this is a high priority based on the number of taught postgraduates mentioning this topic compared to other keywords.

Figure 3.28: Average weighted scores of mentions of 'feedback' (with 95% confidence intervals) across student status

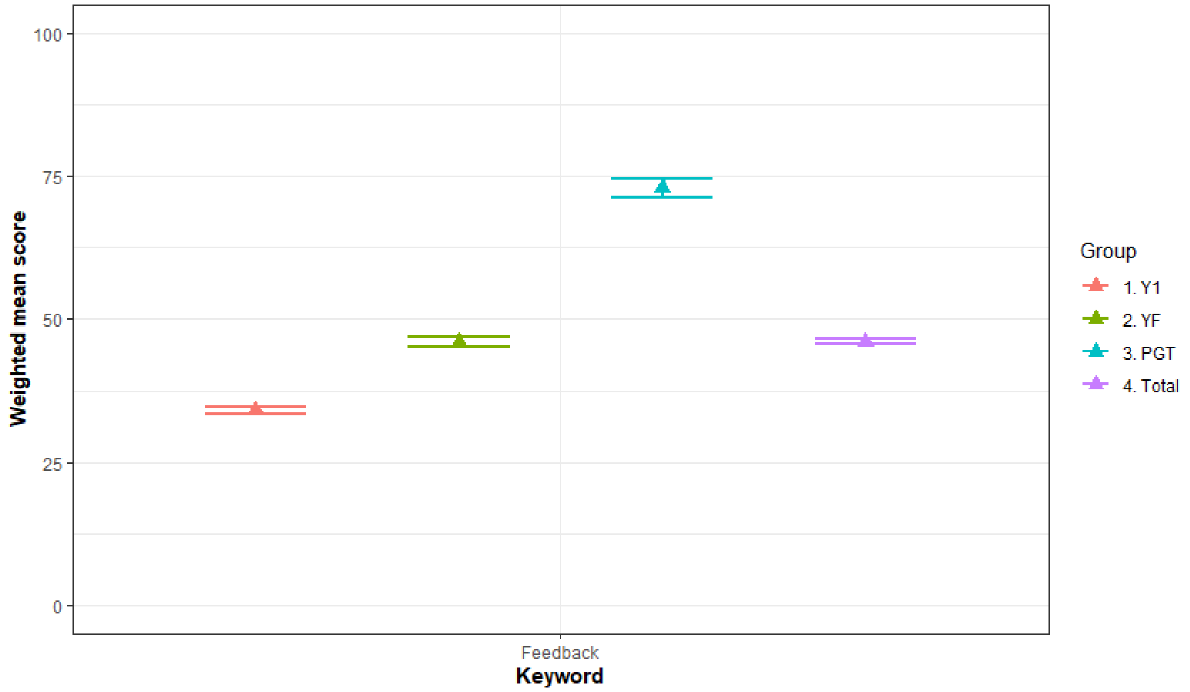
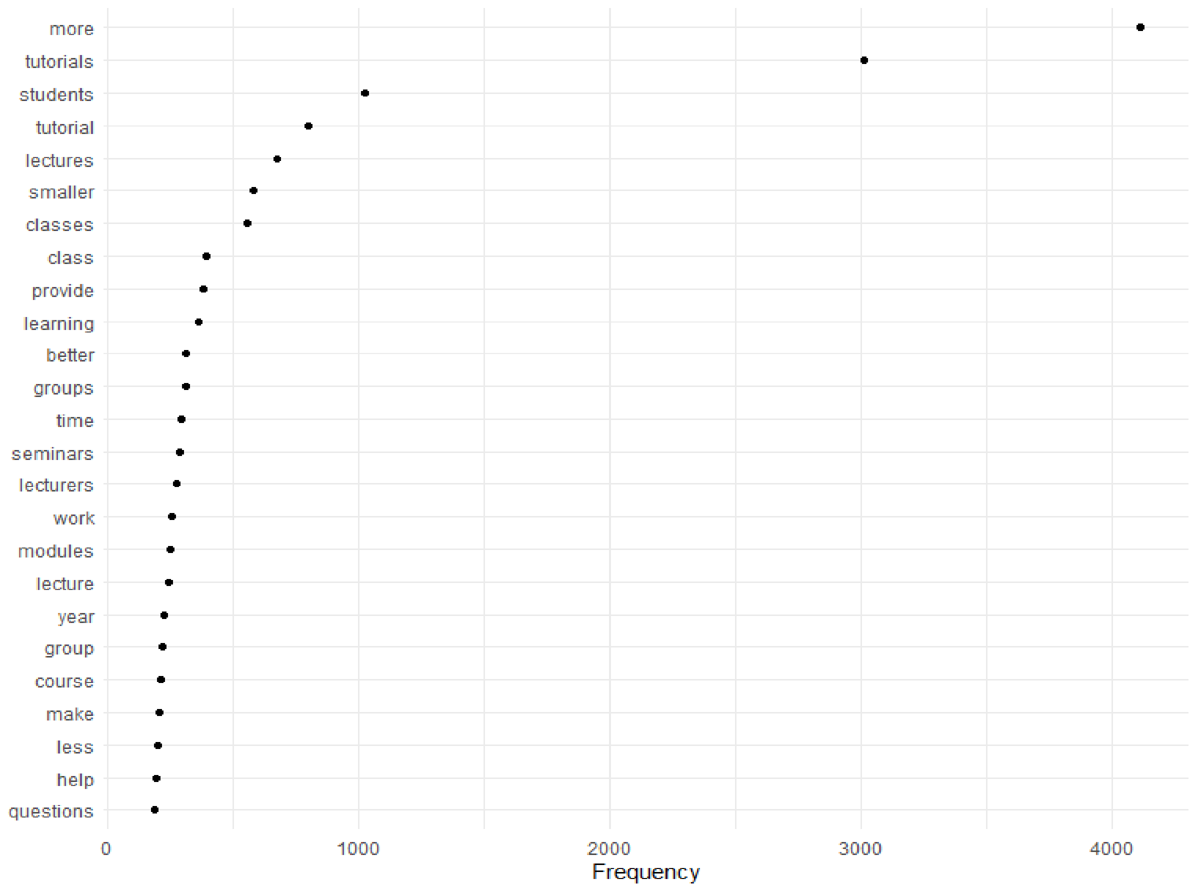
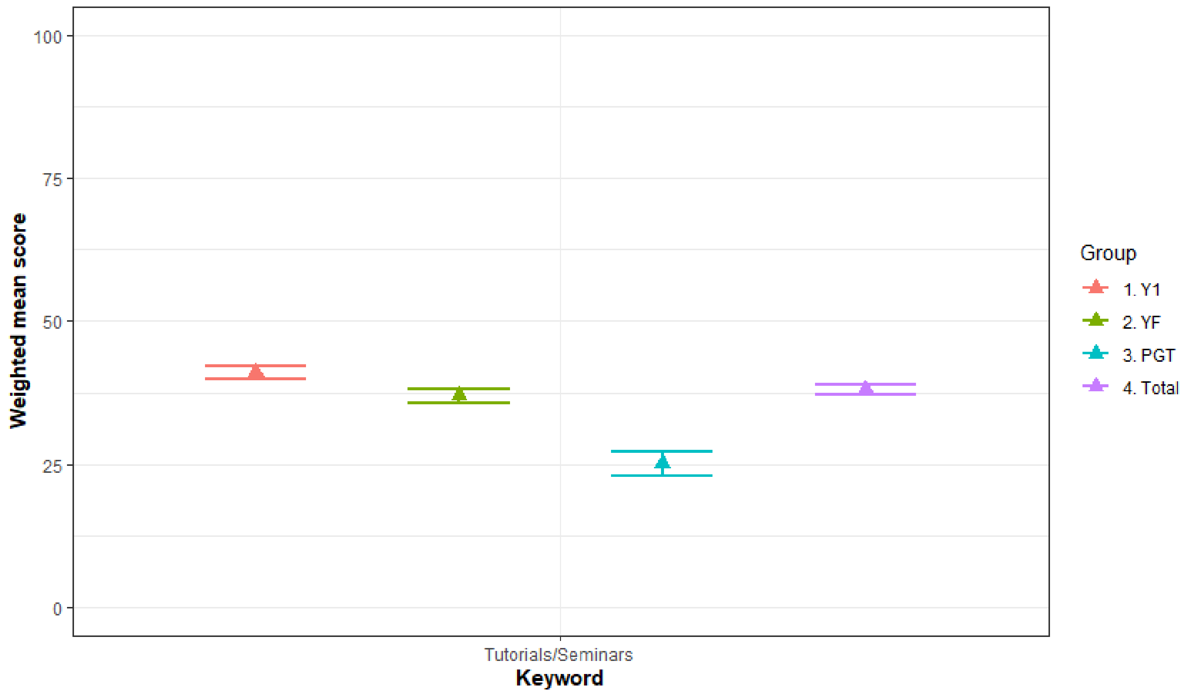


Figure 3.29: Most frequently used words in the 'tutorials/seminars' subset



With regard to the subset of the corpus containing references to tutorials and seminars, Figure 3.29 shows the most frequently used words and again, ‘more’ is the most frequently used word, then ‘tutorials’. Figure 3.30 presents the weighted mean scores and the overall score for all students is around 38 percent. There are statistically significant differences here as taught postgraduates tend to not prioritise this keyword in their mentions and as such, has been downweighted substantially compared to undergraduates which tend to mention this more and want ‘more’ tutorials.

Figure 3.30: Average weighted scores of mentions of ‘tutorials/seminars’ (with 95% confidence intervals) across student status



With regard to the subset of the corpus containing references to help and support, Figure 3.31 shows the most frequently used words and again, ‘more’ is the most frequently used word, ‘help’ and ‘support’ are the third and fourth most used words, respectively. Figure 3.32 presents the weighted mean scores for the proportions and the overall score for all students is around 50. However, this chart shows significant differences across groups as taught postgraduates mention wanting to receive help and support more than their undergraduate counterparts, and this is a high priority based on the number of taught postgraduates mentioning this topic compared to other keywords.

Figure 3.31: Most frequently used words in the 'help/support' subset

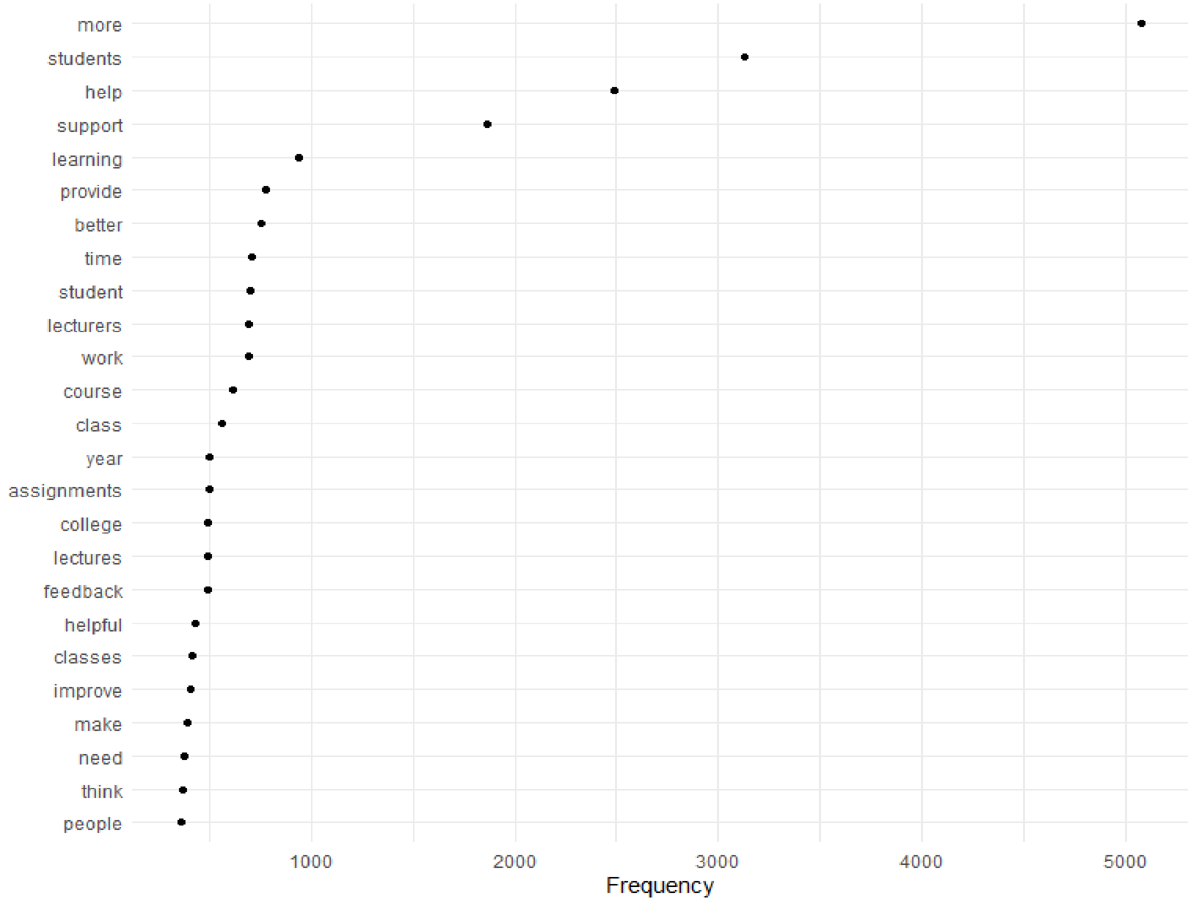


Figure 3.32: Average weighted scores of mentions of 'help/support' (with 95% confidence intervals) across student status

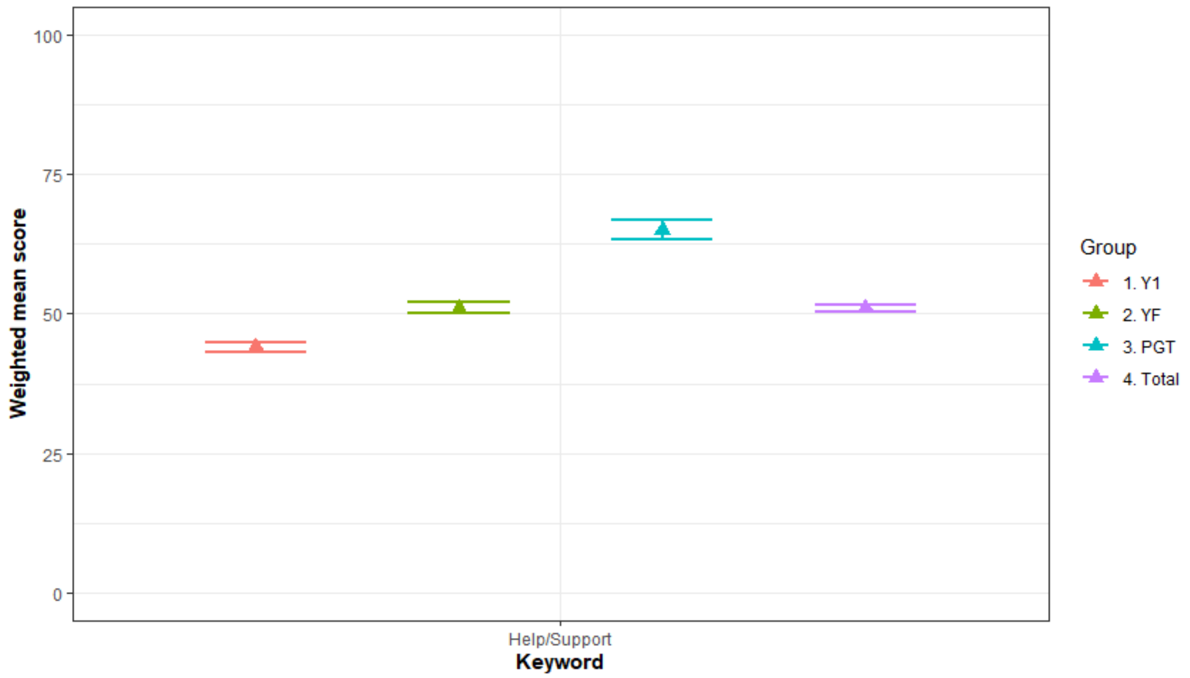
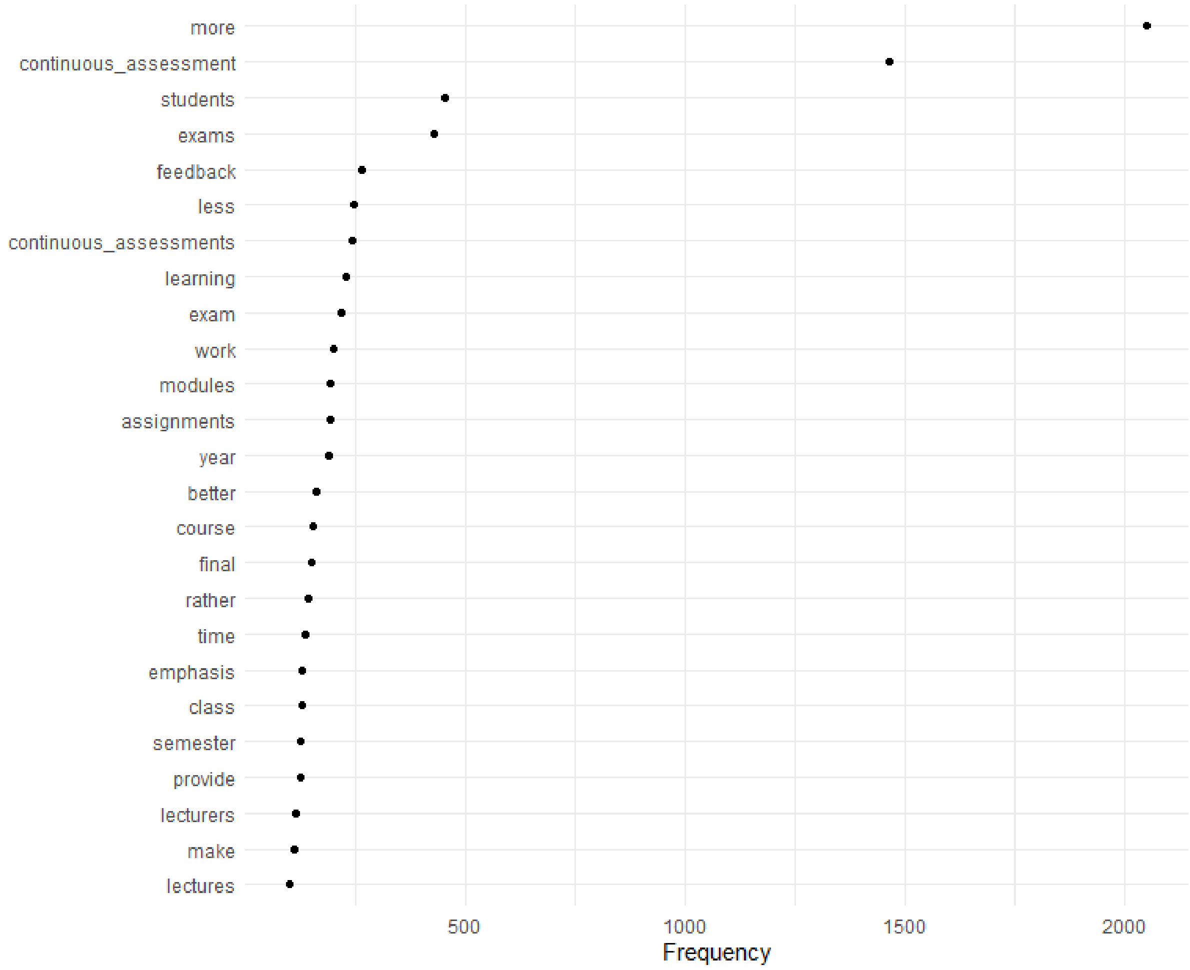
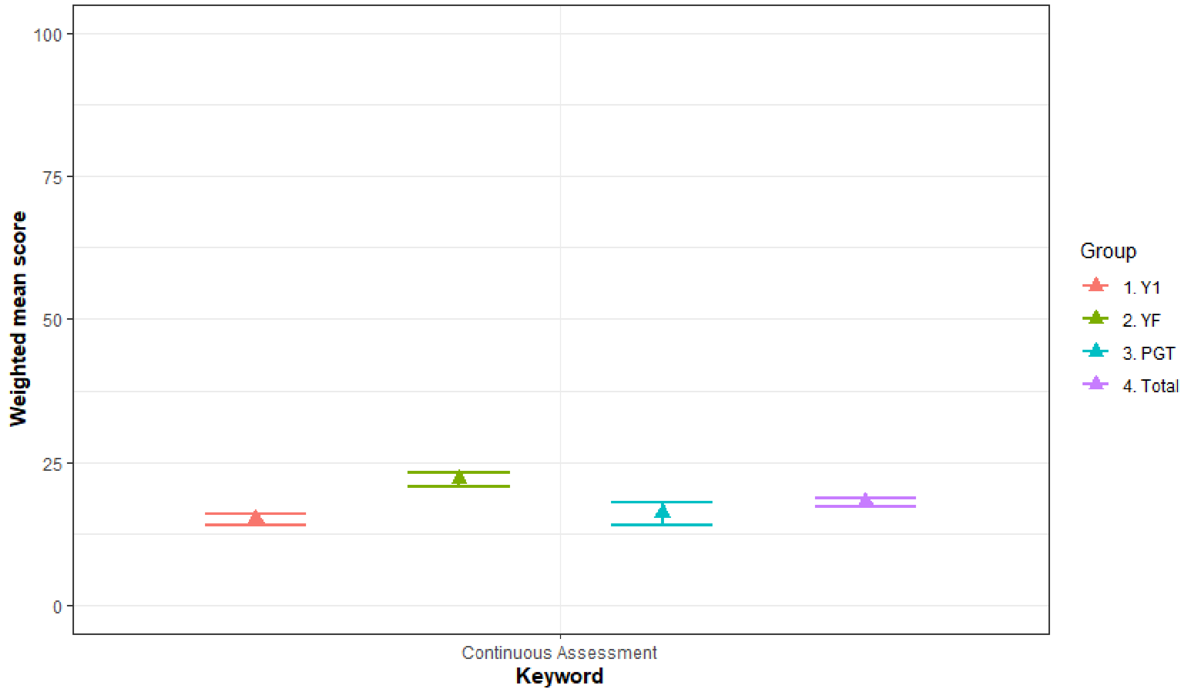


Figure 3.33: Most frequently used words in the ‘continuous assessment’ subset



Finally, for the subset of the corpus containing references to continuous assessment, Figure 3.33 shows the most frequently used words and Figure 3.34 presents the weighted mean scores. Of all the keywords examined in the section, continuous assessment has the lowest scores for all groups. However, this should not be taken as an indication that students do not like it. Rather Figure 3.33 shows that students when mentioning continuous assessment tend to like the idea of having it more than exams, however, in general not many students within groups and overall mention continuous assessment and as such the keyword has been downweighted across the board.

Figure 3.34: Average weighted scores of mentions of 'continuous assessment' (with 95% confidence intervals) across student status

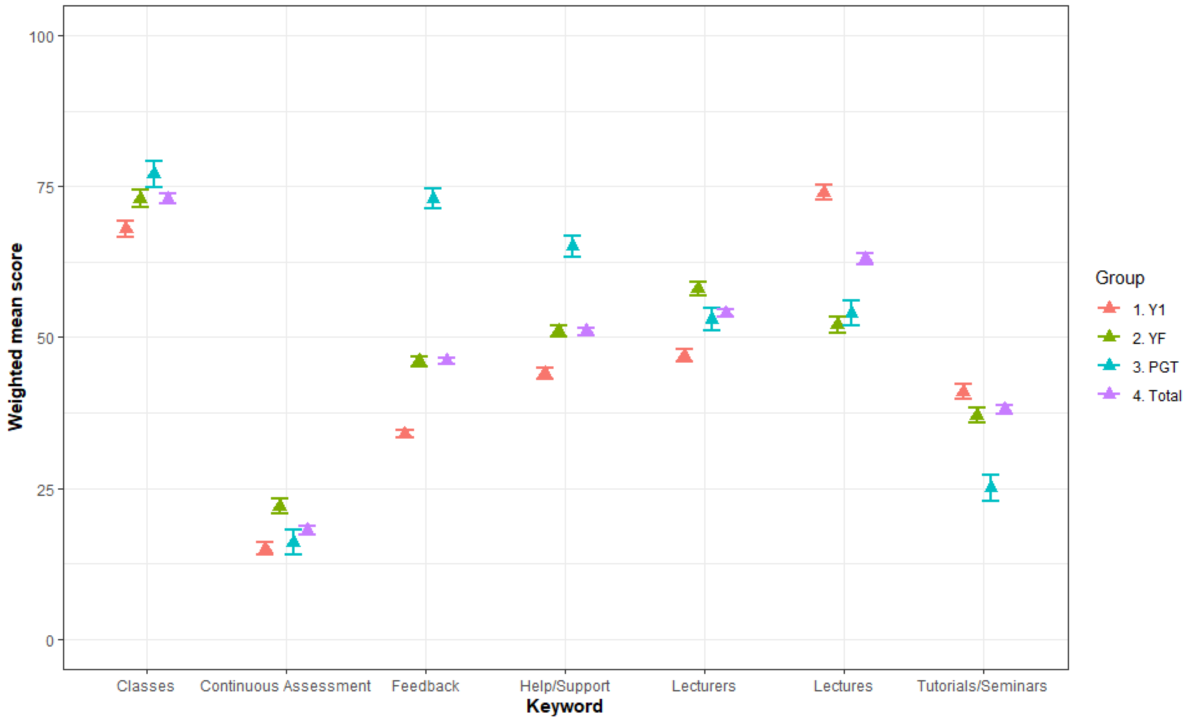


3.6 Conclusions

This chapter has been dominated by the word 'more'. When asked what students would like to see from their institution to improve student engagement, they have typically articulated their thoughts by using the word 'more' as a component of their comment. It has been an inescapable facet of the responses' students provided to this question, but at this point in the analysis we have a coherent idea of what students would like to see 'more' of, taking the number of students making such requests into account.

Figure 3.35 replicates the format of Figure 3.20 though this time presents the weighted mean scores for each keyword and across groups. When presented in this manner how different student group approached providing an answer is readily visible. All students typically would like more contact hours through more classes (with fewer other students in them). Taught postgraduate students desire more feedback on their assignments, whereas other student groups have less demand for this. Taught postgraduate students also would like more support and help throughout their studies. First year students would like more and better lectures and the same for tutorials and seminars. Whereas it appears that there is marginally less demand for this from final year undergraduates and taught postgraduates. Finally, all student groups when they mention continuous assessment typically would like for there to be more of it, and fewer exams, however relative to other keywords it appears that most students have higher priorities in other areas than in continuous assessment.

Figure 3.35: Average weighted scores for all keywords (with 95% confidence intervals) across student status



4. Conclusion

The aim of this study has been to analyse the two qualitative open-text questions contained in the Irish Survey of Student Engagement from 2016 to 2020. These questions are:

> Q1. What does your institution do best to engage students in learning?
>
> Q2. What could your institution do to improve students' engagement in learning?

Q1 asks for an opinion from students and is worded in a manner to obtain a precise response, such as an item or institutional provision that respondents regard as being successful in engaging either themselves or students in general. Q2 poses a more hypothetical query, which asks about a change that could be made, and asks students to weigh up what they have received from their HEI and from this, evaluate what change could be made to improve students' engagement in learning. The structure of each question provided significantly different forms of responses, with responses to Q1 being more wide-ranging and often containing how students felt about or engaged with various items provided by, or part of life in third level education. In contrast, responses to Q2 typically followed a template of a modifier (for example, more/less/better/fewer) along with an item or items.

At the outset and throughout this study, we noted that the main barrier to thorough comprehension of the material was the sheer volume of data provided by students. The comments provided by students would fill 8,000 A4 pages of text with over 1,320,000 words. With this amount of information, it would have been easy to be overwhelmed unless a strategy to break the data down into more manageable units of analysis was adopted.

Our strategy followed a similar path for the analysis for both questions, in that they both broke down the corpus of each question into more comprehensible units first, then reintegrated these units to gain a systematic understanding of students' responses and how these responses provide answers to the questions asked. However, the difference in the how Q1 and Q2 are structured and the responses from students that were evoked meant that the reintegration of units took different directions within the analysis.

The breaking down of each corpus into individual units in the form of words began with translation of responses in Irish into English, a comprehensive review of comments combined with a spellcheck, and compounding of multiword phrases that should be evaluated together. This was followed by entry of the corpus into the statistical software which was then cleaned of punctuation, unnecessary and extraneous words (i.e. stopwords), and tokenisation of the text into individual words based on the whitespace around text characters. The result of all the above was a 'clean' corpus for each question ready for analysis.

This analysis began by examining the overall corpus for each question from 2016 to 2020, and across individual years. This was done by ranking the frequency of words across each year. The result of this was that there appears to be remarkable stability in the corpus over time. Certain words such as 'students', 'learning', 'lecturers', 'lectures', 'class', and 'work' were used consistently more than other words across time. However, this finding is underpinned by the assumption that students across time had similar experiences in their time in higher education. This assumption may not hold in the 2021 data due to the disruption to higher education, and movement from in-person teaching and learning to remote study, caused by the Covid-19 pandemic. As such, the corpus of words used and their frequency in the 2021 data may be substantially different from what has been observed in earlier periods.

The keyness analysis provided us with some evidence that the status of students be it first year undergraduate, final year undergraduate, or taught postgraduate has an influence on their priorities, and these priorities are expressed through the different frequencies of certain words used in their responses, for example, final year students are more likely to mention work experience or placements. However, the keyness analysis also did not show substantive discrepancies across groups, which led us to note that each groups' experience of student life appears to have more in common with each other than their relative differences.

For Q1, the keyness analysis was followed up with latent semantic scaling and this showed that keywords habitually used by students in their responses had a number of positive and negative attributes associated with them. For example, positive words associated with tutorials included 'helpful, 'useful' and 'effective' while negative words associated with the same keyword included 'difficult' and 'difficulty'.

Latent sematic scaling was less useful for Q2 because of the prevalence of a small group of modifier words used to indicate direction and degree. As such, there was less need to capture what students were feeling about keywords through the words used in their responses because there was much less variety in the words used to describe what they would change about these keywords, and the words themselves described the direction and degree to which they would affect items.

For Q1 the analysis was capped off with sentiment analysis of whole student responses which provided us with average sentiment scores for a range of keywords. This provided us with some answers to the overall question asked of students, “What does your institution do best to engage students in learning?”

Overall, it appears that students think that the academic staff of their institution are the best at engaging students in learning. This is followed by the library services, and general facilities provided by their HEI. Student unions, and general means of providing support, help, and assistance also have high mean sentiment scores. The support provided for dissertations is infrequently mentioned but highly placed when it is mentioned. A similar pattern emerges for Academic Learning Centres and Writing Centres, as they are positively evaluated by students, but these are not present in every HEI.

It was also noted that keywords associated with ‘personal’ experiences may have shown high mean sentiment scores due to the positive associations inherent to the keywords themselves. However, this did not mean that HEIs should ignore these facets of students’ experiences. Instead, these may be easier changes to make as they do not require investment and resource allocation within HEIs and could quickly lead to a virtuous cycle reinforcing other aspects of students experiences which would thereby increase student engagement.

Instead of sentiment analysis for Q2, the analysis instead looked at bigrams of words, and showed that responses to Q2 tended to follow a formula of modifier plus item and also showed how effective bigrams were in illustrating typical responses of students. The step beyond this was examining how often modifier words were used against topics that repeatedly recur throughout the corpus. As such, gaining a more complete picture of how students actually answered the question, “What could your institution do to improve students' engagement in learning?”

Student answers to Q2 were dominated by the word ‘more’. When asked what students would like to see from their institution to improve student engagement, they typically articulated their thoughts by using the word ‘more’ as a component of their comment. All students typically would like more contact hours through more classes (with fewer other students in them). This meshes with the low mean sentiment score for ‘small classes’ found in Q1 (cf. p37) as these classes currently have too many students in them to be positively evaluated. Taught postgraduate students desire more feedback on their assignments, whereas other student groups have less demand for this. Taught postgraduate students also would like more support and help throughout their studies. First year students would like more and better lectures, and the same for tutorials and seminars. Whereas it appears that there is marginally less demand for this from final year undergraduates and taught postgraduates. Finally, all student groups when they mention continuous assessment typically would like for there to be more of it, and fewer exams, however relative to other keywords it appears that most students have higher priorities in other areas than in continuous assessment.

In summary, this research project has demonstrated the utility of using computer-assisted analytical techniques to process data which other researchers would have great difficulty in processing and evaluating due to the scale of the material involved. Furthermore, this research has provided a framework for further research to build upon. Our analysis is based upon the open-source statistical software *R* and its graphical interface R Studio, along with a suite of R packages and because of this, all of the techniques utilised within the report are replicable, and our coding framework and code can be extended for further research within the current corpus or further iterations of the Irish Survey of Student Engagement.

Appendix A: Stopwords

Standard Stopwords:

A	About	Above	After	Again	Against	All
Am	An	And	Any	Are	Aren't	As
At	Be	Because	Been	Before	Being	Below
Between	Both	But	By	Cannot	Can't	Could
Couldn't	Did	Didn't	Do	Does	Doesn't	Doing
Don't	Down	During	Each	Few	For	From
Further	Had	Hadn't	Has	Hasn't	Have	Haven't
Having	He	He'd	He'll	Her	Here	Here's
Hers	Herself	He's	Him	Himself	His	How
How's	I	I'd	If	I'll	I'm	In
Into	Is	Isn't	It	It's	Its	Itself
I've	Let's	Me	More	Most	Mustn't	My
Myself	No	Nor	Not	Of	Off	On
Once	Only	Or	Other	Ought	Our	Ours
Ourselves	Out	Over	Own	Same	Shan't	She
She'd	She'll	She's	Should	Shouldn't	So	Some
Such	Than	That	That's	The	Their	Theirs
Them	Themselves	Then	There	There's	These	They
They'd	They'll	They're	They've	This	Those	Through
To	Too	Under	Until	Up	Very	Was
Wasn't	We	We'd	We'll	We're	Were	Weren't
We've	What	What's	When	When's	Where	Where's
Which	While	Who	Whom	Who's	Why	Why's
Will	With	Won't	Would	Wouldn't	You	You'd
You'll	Your	You're	Yours	Yourself	Yourselves	You've

Custom Stopwords:

Lot	Also	One	Us	Get	Can	Lots
Like	???	Its	Just			