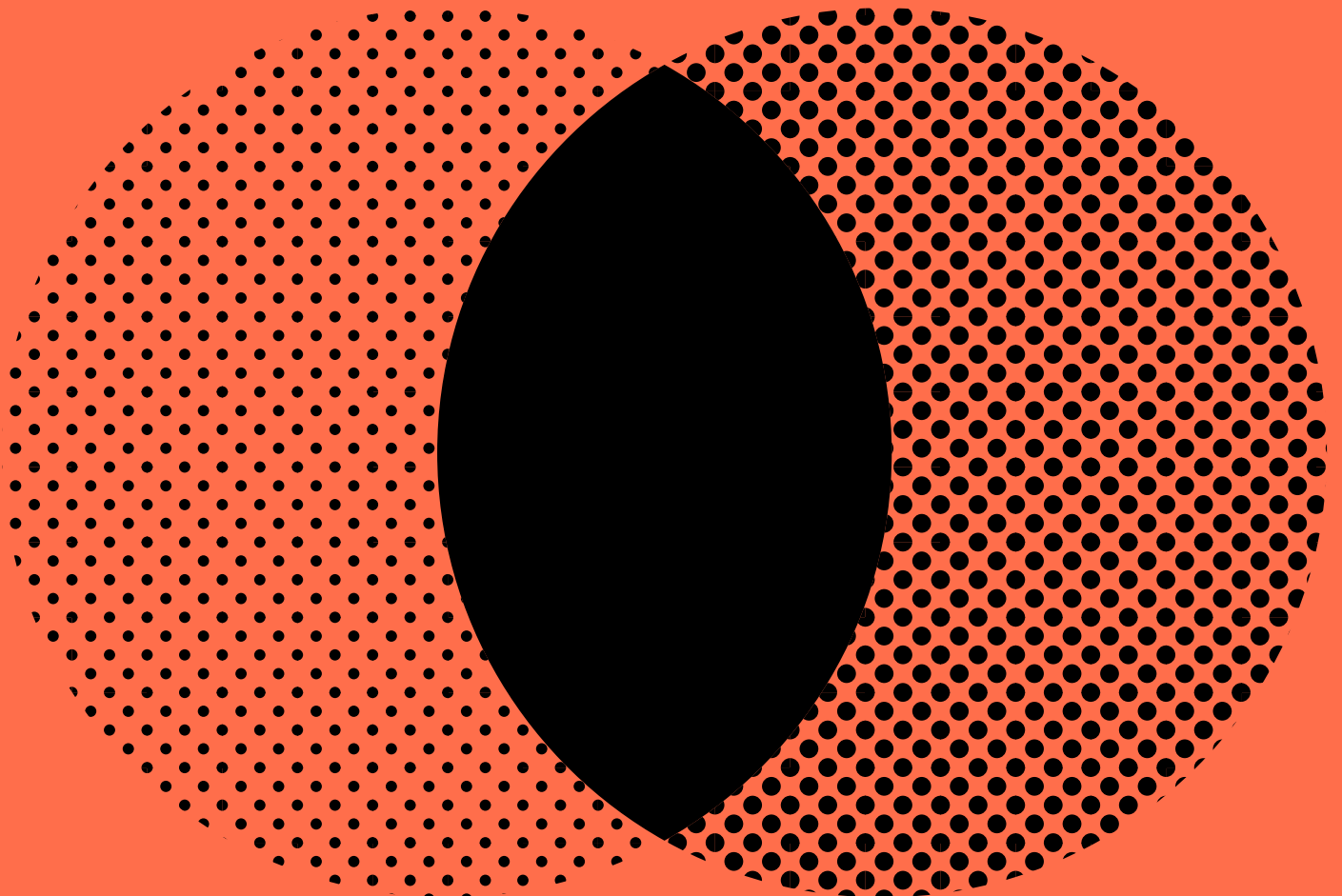


# Irish Survey of Student Engagement

## Results of qualitative data analysis projects

---

Report 3 of 5





## Results of qualitative data analysis projects

**Name of Report:** REVEAL – PGR

**Roadmap to StudentSurvey.ie Qualitative Data Analysis – Postgraduate**

**Authors:** Dr. Seán Lacey

**Organisation/ Institution:** Munster Technological University

### 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.

# REVEAL - PGR

## Roadmap to StudEntSurvey.ie Qualitative Data Analysis – PostGraduate

Seán Lacey, PhD

Department of Mathematics



# MTU

Ollscoil Teicneolaíochta na Mumhan  
Munster Technological University



# PGR Student Survey.ie

May 2021

## Table of Contents

List of Tables .....	4
List of Figures .....	9
Executive Summary.....	11
1. Introduction .....	13
2. Method .....	13
2.1 Overview .....	13
2.2 Data cleaning .....	14
2.3 Data import.....	15
2.3.1 Open-ended questions.....	15
2.3.2 Demographic data.....	15
2.4 Word frequency analysis.....	15
2.5 Text coding and Reference extracting .....	16
2.6 Matrix coding .....	17
2.7 Statistical inference.....	18
3. Results.....	19
3.1 NVivo v. RStudio.....	19
3.2 Research infrastructure and facilities .....	21
3.2.1 Word frequency analysis.....	21
3.2.2 Matrix coding and statistical inference.....	22
3.3 Supervision.....	23
3.3.1 Word frequency analysis.....	23
3.3.2 Matrix coding and statistical inference.....	25
3.4 Research culture .....	26
3.4.1 Word frequency analysis.....	26
3.4.2 Matrix coding and statistical inference.....	28
3.5 Induction, progression arrangements and assessment.....	29
3.5.1 Word frequency analysis.....	29
3.5.2 Matrix coding and statistical inference.....	31
3.6 Development opportunities (including teaching / demonstrating).....	32
3.6.1 Word frequency analysis.....	32
3.6.2 Matrix coding and statistical inference.....	34
3.7 Research skills development.....	35
3.7.1 Word frequency analysis.....	35
3.7.2 Matrix coding and statistical inference.....	36
3.8 Development of other transferable skills .....	37

3.8.1	Word frequency analysis.....	37
3.8.2	Matrix coding and statistical inference.....	39
3.9	Student / staff responsibilities and supports.....	40
3.9.1	Word frequency analysis.....	40
3.9.2	Matrix coding and statistical inference.....	42
3.10	Personal outlook .....	43
3.10.1	Word frequency analysis.....	43
3.10.2	Matrix coding and statistical inference.....	45
3.11	Aspects / elements of research degree programme are most valuable .....	46
3.11.1	Word frequency analysis.....	46
3.11.2	Matrix coding and statistical inference.....	48
3.12	Aspects of research degree experience could be improved.....	49
3.12.1	Word frequency analysis.....	49
3.12.2	Matrix coding and statistical inference.....	51
4.	Summary and Conclusion .....	53
5.	Acknowledgements.....	55
6.	References .....	55
	Appendix RStudio output.....	56
A.1	Research infrastructure and facilities .....	56
A.1.1	Word frequency analysis.....	56
A.1.2	Matrix coding and statistical inference.....	57
A.2	Supervision.....	58
A.2.1	Word frequency analysis.....	58
A.2.2	Matrix coding and statistical inference.....	59
A.3	Research culture .....	60
A.3.1	Word frequency analysis.....	60
A.3.2	Matrix coding and statistical inference.....	62
A.4	Induction, progression arrangements and assessment.....	63
A.4.1	Word frequency analysis.....	63
A.4.2	Matrix coding and statistical inference.....	65
A.5	Development opportunities (including teaching / demonstrating).....	66
A.5.1	Word frequency analysis.....	66
A.5.2	Matrix coding and statistical inference.....	67
A.6	Research skills development.....	68
A.6.1	Word frequency analysis.....	68
A.6.2	Matrix coding and statistical inference.....	70

A.7	Development of other transferable skills .....	71
A.7.1	Word frequency analysis.....	71
A.7.2	Matrix coding and statistical inference.....	72
A.8	Student / staff responsibilities and supports.....	73
A.8.1	Word frequency analysis.....	73
A.8.2	Matrix coding and statistical inference.....	74
A.9	Personal outlook .....	75
A.9.1	Word frequency analysis.....	75
A.9.2	Matrix coding and statistical inference.....	77
A.10	Aspects / elements of research degree programme are most valuable .....	78
A.10.1	Word frequency analysis.....	78
A.10.2	Matrix coding and statistical inference.....	80
A.11	Aspects of research degree experience could be improved.....	81
A.11.1	Word frequency analysis.....	81
A.11.2	Matrix coding and statistical inference.....	82
A.12	Required RStudio packages.....	83

## List of Tables

Table 1	Attributes to NVivo and RStudio .....	14
Table 2	Illustration of the need to explore text coding to determine whether or not word frequency are relevant to the question response – e.g., “What aspects / elements of research degree programme are most valuable?” .....	17
Table 3	Outline of the sample exploration of the references from the open-ended question responses across various types of demographic data included in this project. ....	19
Table 4	Comparison of references outputted between NVivo and RStudio. ....	20
Table 5	NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research infrastructure and facilities”. ....	21
Table 6	Sample of reference extracts from the two most frequently occurring words to the question relating to “Research infrastructure and facilities”. ....	22
Table 7	Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Research infrastructure and facilities”. ....	22
Table 8	Matrix coding of the relationship between HEI type and the two most frequently occurring words to the question relating to “Research infrastructure and facilities”. ....	23
Table 9	NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Supervision”. ....	24
Table 10	Sample of reference extracts from the two most frequently occurring words to the question relating to “Supervision”. ....	24
Table 11	Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Supervision”. ....	25

Table 12 Matrix coding of the relationship between Gender and the two most frequently occurring words to the question relating to "Supervision" .	25
Table 13 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research culture" .	27
Table 14 Sample of reference extracts from the two most frequently occurring words to the question relating to "Research culture" .	27
Table 15 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Research culture" .	28
Table 16 Matrix coding of the relationship between Domicile group and the two most frequently occurring words to the question relating to "Research culture" .	28
Table 17 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Induction, progression arrangements and assessment" .	30
Table 18 Sample of reference extracts from the two most frequently occurring words to the question relating to "Induction, progression arrangements and assessment" .	30
Table 19 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Induction, progression arrangements and assessment" .	31
Table 20 Matrix coding of the relationship between Mode of study and the two most frequently occurring words to the question relating to "Induction, progression arrangements and assessment" .	31
Table 21 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Development opportunities (including teaching / demonstrating)" .	33
Table 22 Sample of reference extracts from the two most frequently occurring words to the question relating to "Development opportunities (including teaching / demonstrating)" .	33
Table 23 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Development opportunities (including teaching / demonstrating)" .	34
Table 24 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to "Development opportunities (including teaching / demonstrating)" .	34
Table 25 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research skills development" .	36
Table 26 Sample of reference extracts from the two most frequently occurring words to the question relating to "Research skills development" .	36
Table 27 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Research skills development" .	36
Table 28 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to "Research skills development" .	37
Table 29 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Development of other transferable skills" .	38

Table 30 Sample of reference extracts from the two most frequently occurring words to the question relating to “Development of other transferable skills” . . . . .	39
Table 31 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Development of other transferable skills” . . . . .	39
Table 32 Matrix coding of the relationship between Domicile group and the two most frequently occurring words to the question relating to “Development of other transferable skills” . . . . .	40
Table 33 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Student / staff responsibilities and supports” . . . . .	41
Table 34 Sample of reference extracts from the two most frequently occurring words to the question relating to “Student / staff responsibilities and supports” . . . . .	41
Table 35 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Student / staff responsibilities and supports” . . . . .	42
Table 36 Matrix coding of the relationship between Mode of study and the two most frequently occurring words to the question relating to “Student / staff responsibilities and supports” . . . . .	42
Table 37 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Personal outlook” . . . . .	43
Table 38 Sample of reference extracts from the two most frequently occurring words to the question relating to “Personal outlook” . . . . .	44
Table 39 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Personal outlook” . . . . .	45
Table 40 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to “Personal outlook” . . . . .	45
Table 41 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Aspects / elements of research degree programme are most valuable” . . . . .	47
Table 42 Sample of reference extracts from the two most frequently occurring words to the question relating to “Aspects / elements of research degree programme are most valuable” . . . . .	48
Table 43 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Aspects / elements of research degree programme are most valuable” . . . . .	48
Table 44 Matrix coding of the relationship between Gender and the two most frequently occurring words to the question relating to “Aspects / elements of research degree programme are most valuable” . . . . .	49
Table 45 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Aspects of research degree experience could be improved” . . . . .	50
Table 46 Sample of reference extracts from the two most frequently occurring words to the question relating to “Aspects of research degree experience could be improved” . . . . .	51
Table 47 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Aspects / elements of research degree experience could be improved” . . . . .	51

Table 48 Matrix coding of the relationship between HEI type and the two most frequently occurring words to the question relating to "Aspects of research degree experience could be improved" . . . . .	51
Table 50 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research infrastructure and facilities". . . . .	57
Table 51 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Research infrastructure and facilities". . . . .	57
Table 52 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Research infrastructure and facilities". . . . .	57
Table 53 Matrix coding of the relationship between HEI type and the two most frequently occurring words to the question relating to "Research infrastructure and facilities". . . . .	58
Table 54 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Supervision". . . . .	59
Table 55 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Supervision". . . . .	59
Table 56 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Supervision". . . . .	59
Table 57 Matrix coding of the relationship between Gender and the two most frequently occurring words to the question relating to "Supervision". . . . .	60
Table 58 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research culture" . . . . .	61
Table 59 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Research culture" . . . . .	61
Table 60 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Research culture". . . . .	62
Table 61 Matrix coding of the relationship between Domicile group and the two most frequently occurring words to the question relating to "Research culture". . . . .	62
Table 62 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Induction, progression arrangements and assessment". . . . .	64
Table 63 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Induction, progression arrangements and assessment" . . . . .	64
Table 64 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Induction, progression arrangements and assessment". . . . .	65
Table 65 Matrix coding of the relationship between Mode of study and the two most frequently occurring words to the question relating to "Induction, progression arrangements and assessment". . . . .	65
Table 66 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Development opportunities (including teaching / demonstrating)". . . . .	67
Table 67 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Development opportunities (including teaching / demonstrating)". . . . .	67

Table 68 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to " Development opportunities (including teaching / demonstrating)" .....	67
Table 69 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to " Development opportunities (including teaching / demonstrating)" .....	68
Table 70 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research skills development" .....	69
Table 71 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Research skills development" .....	70
Table 72 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to " Research skills development" .....	70
Table 73 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to " Research skills development" .....	70
Table 74 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Development of other transferable skills".	72
Table 75 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Development of other transferable skills".	72
Table 76 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Development of other transferable skills".	72
Table 77 Matrix coding of the relationship between Domicile group and the two most frequently occurring words to the question relating to "Development of other transferable skills".	73
Table 78 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Student / staff responsibilities and supports".	74
Table 79 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Student / staff responsibilities and supports".	74
Table 80 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Student / staff responsibilities and supports".	74
Table 81 Matrix coding of the relationship between Mode of study and the two most frequently occurring words to the question relating to "Student / staff responsibilities and supports".	75
Table 82 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Personal outlook".	76
Table 83 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Personal outlook".	77
Table 84 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Personal outlook".	77
Table 85 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to "Personal outlook".	77
Table 86 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Aspects / elements of research degree programme are most valuable" .....	79

Table 87 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Aspects / elements of research degree programme are most valuable". ..... 79

Table 88 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Aspects / elements of research degree programme are most valuable" ..... 80

Table 89 Matrix coding of the relationship between Gender and the two most frequently occurring words to the question relating to "Aspects / elements of research degree programme are most valuable" ..... 80

Table 90 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Aspects of research degree experience could be improved". ..... 82

Table 91 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Aspects of research degree experience could be improved". ..... 82

Table 92 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Aspects of research degree experience could be improved". ..... 82

Table 93 Matrix coding of the relationship between HEI type and the two most frequently occurring words to the question relating to "Aspects of research degree experience could be improved". ..... 83

## List of Figures

Figure 1 Sample references from NVivo relating to responses to "What aspects / elements of research degree programme are most valuable?" ..... 16

Figure 2 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research infrastructure and facilities" ..... 21

Figure 3 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Supervision". ..... 23

Figure 4 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research culture" ..... 26

Figure 5 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Induction, progression arrangements and assessment". ..... 29

Figure 6 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Development opportunities (including teaching / demonstrating)" ..... 32

Figure 7 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research skills development". ..... 35

Figure 8 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Development of other transferable skills" ..... 37

Figure 9 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Student / staff responsibilities and supports". ..... 40

Figure 10 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Personal outlook” ..... 43

Figure 11 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Aspects / elements of research degree programme are most valuable” ..... 46

Figure 12 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Aspects of research degree experience could be improved” ..... 49

Figure 13 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research infrastructure and facilities” ..... 56

Figure 14 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Supervision” ..... 58

Figure 15 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research culture” ..... 60

Figure 16 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Induction, progression arrangements and assessment” ..... 63

Figure 17 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Development opportunities (including teaching / demonstrating)” ..... 66

Figure 18 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research skills development” ..... 68

Figure 19 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Development of other transferable skills” ..... 71

Figure 20 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Student / staff responsibilities and supports” ..... 73

Figure 21 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Personal outlook” ..... 75

Figure 22 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Aspects / elements of research degree programme are most valuable” ..... 78

Figure 23 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Aspects of research degree experience could be improved” ..... 81

## Executive Summary

In this project, a roadmap is presented as a mechanism for qualitative data analysis on PGR StudentSurvey.ie. The roadmap has six steps:

1. Data cleaning;
2. Data import;
3. Word frequency analysis;
4. Text coding and Reference extracting;
5. Matrix coding;
6. Statistical inference.

The structure to REVEAL-PGR is based on the five steps outlined by Feng & Horenstein (2019), however, the additional sixth step – *Statistical inference* – offers a synergy between qualitative analysis with quantitative analysis with results that should complement the narrative around traditional qualitative analysis. The statistical inference step enables the possibility to determine if trends exist in students' responses for a particular year and/or demographic – Higher Education Institution (HEI) type, Gender, Domicile group, Mode of study and Programme type.

The six steps to the roadmap are applied to the 11 open-ended question responses of the PGR StudentSurvey.ie data from 2018 and 2019 using NVivo and RStudio. The rationale to using two software packages is to promote the accessibility of the qualitative analysis for future year's data collection and due to some advantages and disadvantages of each package (see Table 1). The results across packages are consistent with the exception of the responses to *Research infrastructure and facilities* due to how each package manages stemmed variants of a word (see Table 4). Given that the results from the software packages are largely aligned, the results from NVivo are reported in the main body of the report, with the results from RStudio reported in the appendix.

There are two objectives to this project – 1. Qualitative analysis on the 11 open-ended questions from 2018 and 2019 data; 2. Introduce a roadmap to promote the accessibility of the analysis of the unstructured qualitative data emerging from PGR StudentSurvey.ie. From the analysis of the 11 open-ended questions the main findings are:

1. *Research infrastructure and facilities* – The two most frequent occurring words are *access* and *funding*, with the former being more prevalent in Institutes of Technology and latter in Universities and Other Institutions;
2. *Supervision* – Responses varied from supervisors being excellent to supervisors requiring training in supervision;
3. *Research culture* – Suggests a lack of collaboration between departments and some students;
4. *Induction, progression arrangements and assessment* – Students predominantly respond with there being a lack of induction training;
5. *Development opportunities (including teaching / demonstrating)* – Responses to this question are mixed between teaching being a great experience to there being no opportunity to avail of teaching/tutoring;

6. *Research skills development* – Students’ responses to this question varied from a lack of development and training available to training being excellent;
7. *Development of other transferable skills* – Responses varied from supervisor being a great support in developing skills, to there being no opportunity and students being left to acquire skills independently;
8. *Student / staff responsibilities and supports* – Responses range from a positive experience with regards supports to none being existent;
9. *Personal outlook* – A lack of support to students’ is the most prevalent response from students to this question;
10. *Aspects / elements of research degree programme are most valuable* – Supervisors are the most prevalent valuable aspect from the analysis on this question’s responses;
11. *Aspects of research degree experience could be improved* – Similar to responses to some previous open-ended questions, the need for additional student support is most prevalent when student’s reply to this question.

From applying steps 5 and 6 to the roadmap, it is apparent that student’s response to the 11 open-ended questions respectively are consistent across years of data collection. When reviewing responses collectively across all 11 open-ended questions it would be appear that supervisors are the main positive influence on a student’s research career, while additional student support would appear to be the most prevalent aspect that needs attention. It will be interesting to see if the findings from this project are similar to data collected in 2021 given COVID-19.

Overall, this project showcases the accessibility of a roadmap to national PGR StudentSurvey.ie data through analysis of the 11 open-ended question response from 2018 and 2019. However, it will potentially be at respective HEI level, due to the diverse nature to HEI’s in Ireland, where the impact and usefulness of **REVEAL – PGR (Roadmap to StudEntSurVey.ie Qualitative Data AnaLysis – PostGRaduate)** may be realised.

## 1. Introduction

The aim of this project is to provide a roadmap that will facilitate future analysis to the responses of the open ended questions from the PGR StudentSurvey.ie. The roadmap should enable promotion of the accessibility of results from the qualitative analysis of the data rich information that can be obtained from the responses to the PGR StudentSurvey.ie open-ended questions. This project will demonstrate the accessibility of the roadmap through analysing the unstructured qualitative data emerging from PGR StudentSurvey.ie since 2018.

This project will use two packages to promote and encourage the future application of the roadmap provided – i.e., NVivo and RStudio.

- NVivo is a qualitative data analysis computer software package produced by QSR International. NVivo helps qualitative researchers to organise, analyse and find insights in unstructured or qualitative data like interviews, open-ended survey responses, journal articles, social media and web content (<https://www.qsrinternational.com/>).
- RStudio is a free and open source package that supports direct code execution, as well as tools for plotting, history, debugging and workspace management (<https://rstudio.com/>).

The two packages have various advantages and disadvantages, to which a sample of these are summarised in Table 1. In order to promote the easy access to this project's roadmap, the aim of the project is to perform the qualitative data analysis using both packages, so that PGR StudentSurvey.ie can promote and showcase both methods of the roadmap to satisfy respective Higher Education Institutions (HEIs) individual requirements.

## 2. Method

### 2.1 Overview

Feng & Horenstein (2019) have outlined a successful five-step process to analyse open-ended responses to a survey:

1. Data cleaning;
2. Data import;
3. Word frequency analysis;
4. Text coding and Reference extracting;
5. Matrix coding.

From this approach Feng & Horenstein (2019) projected a deeper insight into the participants' experiences and perspectives about the Clinical Translational Science Institute programs. The methodology described in this paper is applicable to other educational or program evaluations - e.g., PGR StudentSurvey.ie. Although Feng & Horenstein (2019) used NVivo only in their research, RStudio has also been found to be impactful with qualitative data analysis (Estrada 2017) and both software packages are used in this project.

Attributes	NVivo	RStudio
Open source	X (€599 for a single user academic licence)	✓ (Open source)
Environment	✓ (“Point and Click”)	X (Programming environment)
Data types*	✓ (Allows for all data types – i.e., text, images, videos, etc.)	X (Only accepts numeric and text data)
Analysis	X (Performs qualitative analysis only, hence, any requirement to link qualitative codes to quantitative data and/or complement with statistical inference will require a different software package)	✓ [Allows for storage and analysis of all data (qualitative and quantitative) within the same project, hence, allowing for accessible link between qualitative codes and quantitative data (if required)]
Promotes Open Science	✓ (Project function)	✓ (Script from the programming)

Table 1 Attributes to NVivo and RStudio.

\* Not of immediate relevance to the current project.

This project will use the Feng & Horenstein (2019) five-step process, with an additional sixth step of statistical inference, and apply to the PGR StudentSurvey.ie data.

## 2.2 Data cleaning

This step is has two parts:

1. Removing any identifiable information: StudentSurvey.ie carried out this process before sharing the data. However, text related to StudentSurvey.ie’s data cleaning was removed before importing into the appropriate software package - e.g., <coursetitle removed>, <jobtitle removed>, <name removed>, <otheridentifier removed>, etc.;
2. Translating from Irish to English: Given that StudentSurvey.ie is carried out in Ireland with some native Irish speakers included in the population, some of the student responses needed to be translated to English before importing. From 5,704 student responses, 13 (0.2%) students, all from 2019, responded in Irish;

Addressing misspellings in the raw data is not a requirement for this analysis as the word frequency analysis outlined in Section 2.4 includes stemmed variants for words and hence adjusts for misspellings.

## 2.3 Data import

This step involves importing the open-ended responses and respondent's demographic data.

### 2.3.1 Open-ended questions

1. Research infrastructure and facilities,
2. Supervision,
3. Research culture,
4. Induction, progression arrangements and assessment,
5. Development opportunities (including teaching / demonstrating),
6. Research skills development,
7. Development of other transferable skills,
8. Student / staff responsibilities and supports,
9. Personal outlook,
10. Aspects / elements of research degree programme are most valuable,
11. Aspects of research degree experience could be improved.

### 2.3.2 Demographic data

1. Year of fieldwork (2018 and 2019),
2. HEI type ( Institutes of Technology, Universities and Other Institutions),
3. Gender (Female, Male),
4. Domicile group (Irish, non-Irish),
5. Mode of study (Full-time, Part-time/Remote),
6. Programme type (Masters, PhD).

## 2.4 Word frequency analysis

Carley (1993) showed the approach of word frequency analysis as beneficial due to important and significant words being used more frequently in open-ended question responses. Furthermore, Jackson & Trochim (2002) stated that word count has its advantages in identifying patterns more easily, verifying a hypothesis and maintaining analytic integrity.

Words included in this analysis have five or more letters (for example, words like “we”, “the”, “is”, etc. are excluded). This part of the analysis has three steps:

1. The open-ended responses of each question are run through the word frequency analysis using words with stemmed variants (e.g., “research, researches, researcher, researchers” are counted as “research”) with their weighted percentages calculated. Determination of the stemmed variant is the main difference between NVivo and RStudio. The former uses the stemmed variant of the most frequent version of the word, while the latter reduces the word down to its core – e.g., *funding* is reduced to *fund* in RStudio. It is with the reference of *funding* in response to *Research infrastructure and facilities* that the two packages have one of two

conflicts in this project. In NVivo, the reference *funding* is picked up and is in fact the most frequent used word, see Figure 2 and Table 5, for responses to *Research infrastructure and facilities*. While in RStudio *funding* is reduced to *fund* and with the word frequency analysis only including words of 5 or more letters, *fund* is not referenced, see Figure 13 and Table 49. Of course, the minimum word length could be reduced, but this has not been applied in this project simply to highlight a subtle difference between NVivo and RStudio;

2. Feng & Horenstein (2019) recommend the most frequently occurring words, with a weighted percentage of at least 0.5%, to serve as coded words or references (labelled as nodes in NVivo);
3. The visual for the word frequency analysis is a word cloud.

## 2.5 Text coding and Reference extracting

The most frequently occurring words (with the stemmed variants) are coded as response content – i.e., references. The reference relating to the frequently occurring word is used to extract the response content. For example, in Figure 1 the stemmed variant of *research* is the most frequently occurring word to the question *What aspects / elements of research degree programme are most valuable?*. The references associated with the word are varied – e.g., *Training and publication of **research**, Collaboration with peers in the same **research** field*, etc. From the references extracted it is possible to develop themes to question responses, but this would be very much dependent on a research hypothesis and is an example of how the roadmap outlined in this project could be expanded in the future.

### Reference 1 - 0.01% Coverage

Training and publication of **research**

### Reference 2 - 0.01% Coverage

Collaboration with peers in the same **research** field.

### Reference 3 - 0.01% Coverage

Relative freedom to pursue chosen **research**. Academic excellence throughout the Department/University Extra-curricular lectures complimentary to my chosen field Inclusivity

### Reference 4 - 0.01% Coverage

Supervision; training; other PhD students and staff; atmosphere (regarding **research** and otherwise)

Figure 1 Sample references from NVivo relating to responses to “What aspects / elements of research degree programme are most valuable?”

For the purposes of this project five examples of the two most frequently occurring words will be presented to complement the word frequency analysis. Additional information, if required, can be extracted from a user’s NVivo project file and/or RStudio script.

When reviewing a reference, a user could find that a frequently occurring word is not contributing to the answer of the question. For example, when students are asked *What aspects / elements of research degree programme are most valuable?* the most frequent word to appear in a response is *research* (with the stemmed variant), see Figure 1 and Table 2. However, this word is not contributing to the answer of the question. For example, a user would expect a stemmed variant of the word *research* to be mentioned in response to a question about research – e.g., *Training and publication of research*, *Collaboration with peers in the same research field*, etc. But *research* would not be the most accurate theme to extract from the question. For example, if the response is *Training and publication of research*, then *training* and/or *publication* could be a more appropriate theme in relation to the question *What aspects / elements of research degree programme are most valuable?*. However, these would be masked by the frequency of the stemmed variant *research*. If a word is not contributing to the response to a question, it can be “blocked” from the word frequency analysis, with the word frequency analysis step repeated and subsequent text coding updated. For this project the stemmed variant of *research* was blocked from the word frequency analysis for all 11 open-ended questions. Most of the questions used the word *research* in the text, hence the stemmed variant of *research* appears in the response. Note that blocking the stemmed variant of *research* does not mean that any response with the word *research* is excluded, but more so that the stemmed variant of *research* will not appear in the word frequency analysis, word cloud and subsequent analysis.

Word	Length	Count	Weighted Percentage (%)	Similar Words
<b>Research</b>	<b>8</b>	<b>1063</b>	<b>5.23</b>	<b>research, researched, researcher, researchers, researches, researching</b>
supervisors'	12	648	3.19	supervisor, supervisors, supervisors'
Skills	6	520	2.56	skill, skilled, skilling, skills
Support	7	410	2.02	support, supported, supporting, supportive, supports
Opportunity	11	334	1.64	opportunities, opportunity

Table 2 Illustration of the need to explore text coding to determine whether or not word frequency are relevant to the question response – e.g., “What aspects / elements of research degree programme are most valuable?”

Deciding to add a word to a blocked list of words can be subjective, hence why there can be such variability in results from qualitative analysis. (The accompanying video to the final project will outline this process for both NVivo and RStudio packages.)

## 2.6 Matrix coding

The use of the matrix coding feature allows a user to make comparisons and apply an exploratory narrative (as opposed to confirmatory narrative) on responses to open-ended questions with demographic data (Year of fieldwork, HEI type, Gender, Domicile group, Mode of study, Programme

type). Essentially, the matrix coding will enable the exploration of trends in references across years and/or demographics and/or with respect to responses to the closed ended questions. (The latter being beyond the scope of this project.) For example, are references extracted from a particular question response associated with a particular year and/or particular type of student? The matrix coding enables this type of exploration.

It is within the matrix coding that the second main difference between NVivo and RStudio exists. NVivo can provide the matrix coding on the stemmed variants of a reference, however RStudio does not do this. Hence, the counts in the respective matrices (also called contingency tables or crosstabulations) will differ between the packages. Although the statistical inference (step 6) results align.

## 2.7 Statistical inference

From the matrix produced in step 5, it is possible to apply formal statistical testing to enable reliable conclusions to be drawn from the matrix output. The matrix coding enables the exploration of trends in references across years and/or demographics and/or with respect to responses to the closed ended questions, but the statistical inference will allow conclusions to be drawn in relation to the trend visible in the matrix being statistically significant – i.e., reliable – or not statistically significant – i.e., occurred due to chance. The testing will be possible within RStudio, but Excel is used to draw conclusions from the matrix coding produced in NVivo. Statistical inference is carried out over two steps:

1. Year - Initially a yearly comparison is made to determine if responses have varied from 2018 to 2019. A chi-squared test for association is applied here. To demonstrate how the test can be used, the association between *Year of fieldwork* and the two most frequently occurring words in a response to a question are tested. All statistical test results are interpreted using a 5% level of significance.
2. Demographic data - If there is a not statistically significant difference in responses from 2018 to 2019, then when testing relationships between students responses to the open-ended questions with respect to demographic data, year does not need to be controlled for – i.e., analysis does not need to be split by year. While if a statistically significant difference in responses from 2018 to 2019 exists, then when testing relationships – i.e., investigating trends - between students responses to the open-ended questions with respect to demographic data, year will need to be controlled for – i.e., analysis would need to be split by year.

There is a wide range of possibilities in how to explore for trends in references for a particular year, particular demographic and/or how students responded to the closed-ended questions. For example, if the association between all 6 demographic variables - Year of fieldwork, HEI type, Gender, Domicile group, Mode of study and Programme type - with each open-ended question was carried out, this would lead to at least 66 matrices. The rationale for applying step 5 and 6 in this project is to show the depth to the potential analysis available. The level of detail to the matrix coding could be dependent on a particular HEI. For example, a particular HEI could be interested in how Non-Irish PhD

students responded to the open-ended question relating to *Supervision* in 2019. The matrix coding and statistical inference steps enables this type of exploration.

For the purposes of this project a sample of the type of analysis available is summarised in Table 3, with comparisons made in relation to the two most frequently occurring words from the open-ended question responses across various types of demographic data.

Open-ended questions	Demographic data
Research infrastructure and facilities	HEI type
Supervision	Gender
Research culture	Domicile group
Induction, progression arrangements and assessment	Mode of study
Development opportunities (including teaching / demonstrating)	Programme type
Research skills development	Programme type
Development of other transferable skills	Domicile group
Student / staff responsibilities and supports	Mode of study
Personal outlook	Programme type
Aspects / elements of research degree programme are most valuable	Gender
Aspects of research degree experience could be improved	HEI type

Table 3 Outline of the sample exploration of the references from the open-ended question responses across various types of demographic data included in this project.

### 3. Results

#### 3.1 NVivo v. RStudio

When carrying out the word frequency analysis in the two software packages, NVivo uses the stemmed variant of the most frequent version of the word, while the RStudio reduces the word down to its core – e.g., *funding* is reduced to *fund* for the responses to *Research infrastructure and facilities*. In NVivo, the reference *funding* is picked up and is the most frequent used word, see Figure 2 and Table 5, for *Research infrastructure and facilities* While in RStudio *funding* is reduced to *fund* and with the word frequency analysis only including words of 5 or more letters, *fund* is not referenced, see Figure 13 and Table 49. Of course, the minimum word length could be reduced, but this has not been applied in this project simply to highlight a subtle difference between NVivo and RStudio.

Open-ended questions	NVivo v. RStudio (references presented)
Research infrastructure and facilities	Minor inconsistency*
Supervision	Consistent
Research culture	Consistent
Induction, progression arrangements and assessment	Consistent
Development opportunities (including teaching / demonstrating)	Consistent
Research skills development	Consistent
Development of other transferable skills	Consistent
Student / staff responsibilities and supports	Consistent
Personal outlook	Consistent
Aspects / elements of research degree programme are most valuable	Consistent
Aspects of research degree experience could be improved	Consistent

Table 4 Comparison of references outputted between NVivo and RStudio.

\* Due to software packages difference in handling stemmed variants.

It is within the matrix coding that the second main difference between NVivo and RStudio exists. NVivo can provide the matrix coding on the stemmed variants of a reference, however RStudio does not do this. Hence, the counts in the respective matrices differ between the packages, although the statistical inference results are aligned.

Given that the results from the software packages are largely aligned, the results from NVivo are reported in this section, with the results from RStudio reported in the appendix.

### 3.2 Research infrastructure and facilities

#### 3.2.1 Word frequency analysis



Figure 2 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research infrastructure and facilities”.

From Figure 2 and Table 5 it is apparent that the two most frequent occurring words from the question relating to *Research infrastructure and facilities* are *funding* and *access/space*. Table 6 provides a sample of references extracted from students highlighting a lack of funding to attend workshops, conferences, and maintenance of equipment as being a difficulty, while students highlight access to facilities, software packages and journals as an issue under *Research infrastructure and facilities*.

Word	Length	Count	Weighted Percentage (%)	Similar Words
Funding	7	275	1.44	funded, funding, funding', funds
Access	6	253	1.32	access, accessed, accessibility, accessible, accessing
Space	5	253	1.32	space, space', spaces
Library	7	217	1.13	libraries, library
Facility	8	194	1.01	facilities, facility
available	9	141	0.74	avail, availability, available, availed, available, availing
computing	9	121	0.63	computation, computational, computationally, computer, computers, computing
Cover	5	118	0.62	cover, covered, covering, covers
university	10	117	0.61	universal, universities, university
equipment	9	115	0.60	equipment, equipments, equipped
Office	6	106	0.55	office, officer, officers, offices
Stipend	7	106	0.55	stipend, stipends

Table 5 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research infrastructure and facilities”.

<b>Funding</b>
<ul style="list-style-type: none"> <li>• The lack of funding makes it very difficult to attend workshops or conferences highly relevant to my research.</li> <li>• The facilities available to research students in *** are excellent. However, there needs to be funding put aside for the maintenance of equipment.</li> <li>• Research would not be possible without the funding.</li> <li>• Travel/accommodation to at least one conference funded by the college would be nice.</li> <li>• Funding to buy additional equipment that arises as an account of the ongoing research would be very welcome.</li> </ul>
<b>Access</b>
<ul style="list-style-type: none"> <li>• I had access to all of the facilities required to complete my research but few of them are on campus.</li> <li>• dont have access to all requisite journals eg taylor and francis.</li> <li>• More access to desks for masters students in some wings, fix microwaves, have communal areas cleaned better.</li> <li>• Access to NVivo is required to analyse the data which isn't readily available to me.</li> <li>• More journal access.</li> </ul>

Table 6 Sample of reference extracts from the two most frequently occurring words to the question relating to "Research infrastructure and facilities".

### 3.2.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Research infrastructure and facilities* from 2018 to 2019 and HEI type are explored, Table 7 and Table 8, respectively, to determine where or not trends exist in how student's respond to the question.

Year of fieldwork	Research infrastructure and facilities		Total
	Access	Funding	
2018	138 (50.2%)	137 (49.8%)	275 (100.0%)
2019	115 (45.5%)	138 (54.5%)	253 (100.0%)
Total	253 (47.9%)	275 (52.1%)	528 (100.0%)

Table 7 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Research infrastructure and facilities".

For *Research infrastructure and facilities*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Access* and *Funding* is not statistically significant ( $p = 0.2774$ ). This means that it is not one year in particular that issues with access and funding are highlighted.

HEI type	Research infrastructure and facilities		Total
	Access	Funding	
Institutes of Technology	61 (59.8%)	41 (40.2%)	102 (100.0%)
Universities	186 (45.5%)	223 (54.5%)	409 (100.0%)
Other Institutions	6 (35.3%)	11 (64.7%)	17 (100.0%)
<b>Total</b>	<b>253 (47.9%)</b>	<b>275 (52.1%)</b>	<b>528 (100.0%)</b>

Table 8 Matrix coding of the relationship between HEI type and the two most frequently occurring words to the question relating to "Research infrastructure and facilities".

From the two most frequently occurring words, the majority of references from Institutes of Technology were related to *access*, while the majority of references from Universities and Other Institutions is funding ( $p = 0.0199$ ).

### 3.3 Supervision

#### 3.3.1 Word frequency analysis



Figure 3 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Supervision".

From Figure 3 and Table 9 it is apparent that the two most frequent occurring words from the question relating to *Supervision* are *supervisor* and *students*. Table 10 provides a sample of references extracted from students providing praise upon supervisor support to opinions that some supervisors require training on supervision and management.

Word	Length	Count	Weighted Percentage (%)	Similar Words
supervisor	10	859	7.82	supervisor, supervisors, supervisors'
students	8	217	1.97	student, student', students, students'
supervision	11	185	1.68	supervise, supervised, supervisor, supervising, supervision, supervision', supervisions
support	7	165	1.50	support, supported, supporting, supportive, supports
excellent	9	80	0.73	excellent
training	8	77	0.70	trained, training, trainings
meetings	8	72	0.66	meeting, meetings
provide	7	67	0.61	provide, provided, provides, providing
helps	5	66	0.60	helped, helpful, helping, helps
feedback	8	65	0.59	feedback, feedbacks
contact	7	58	0.53	contact, contactable, contacted, contacting, contacts
project	7	55	0.50	project, projects

Table 9 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Supervision".

Supervisor
<ul style="list-style-type: none"> <li>• There needs to be supervisor training.</li> <li>• I am very pleased with my supervisor.</li> <li>• I could not have done as well as I did throughout the PhD without the support of my supervisor. I am very grateful.</li> <li>• I have an excellent supervisor.</li> <li>• Supervisors very generous with their time, contacts, knowledge.</li> </ul>
Students
<ul style="list-style-type: none"> <li>• Even though I appreciate the support of my supervisors, the lead supervisor always wants to force her opinion on the student.</li> <li>• The main method of retrieving information on *** courses is via informal discussions with fellow PhD students. There should be more collaboration of social events for PhD students in an aim to share information.</li> <li>• Supervisors are not monitored for the way they treat students, with some pushing political and religious agendas and putting students in an impossible position. Emotional abuse is not uncommon for certain supervisors.</li> <li>• Supervisors seem to get training about teaching but they also need management training for their research students as this seems to be a common source of problems, mismanagement.</li> <li>• Support for research students is particularly lacking in the summer period.</li> </ul>

Table 10 Sample of reference extracts from the two most frequently occurring words to the question relating to "Supervision".

### 3.3.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Supervision* from 2018 to 2019 and gender are explored, Table 11 and Table 12, respectively, to determine where or not trends exist in how student's respond to the question.

Year of fieldwork	Supervision		Total
	Students	Supervisor	
2018	134 (21.2%)	499 (78.8%)	633 (100.0%)
2019	83 (18.7%)	360 (81.3%)	443 (100.0%)
Total	217 (20.2%)	859 (79.8%)	1076 (100.0%)

Table 11 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Supervision".

For *Supervision*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Students* and *Supervisor* is not statistically significant ( $p = 0.3276$ ). This means that it is not one year in particular that most frequently occurring words from the word frequency analysis are prevalent.

Gender	Supervision		Total
	Students	Supervisor	
Female	145 (20.3%)	569 (79.7%)	714 (100.0%)
Male	72 (19.9%)	290 (80.1%)	362 (100.0%)
Total	217 (20.2%)	859 (79.8%)	1076 (100.0%)

Table 12 Matrix coding of the relationship between Gender and the two most frequently occurring words to the question relating to "Supervision".

For *Supervision* and the two most frequently occurring words – *Students* and *Supervisor* – the difference in distribution of references between males and females is not statistically significant ( $p = 0.8715$ ) – i.e., the two most frequently occurring words to the question relating to *Supervision* are consistent for both males and females.

### 3.4 Research culture

#### 3.4.1 Word frequency analysis



Figure 4 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research culture”.

From Figure 4 and Table 13 it is apparent that the two most frequently occurring words from the question relating to *Research culture* are *students* and *departments*. Table 14 provides a sample of references extracted from students highlighting that there are instances “good collaboration” among students to a direct opposite of there being a “toxic...atmosphere”. While in terms of *departments*, students’ feedback appears to stem around a lack of collaboration between departments.

Word	Length	Count	Weighted Percentage (%)	Similar Words
students	8	320	3.40	student, students, students'
departments	11	272	2.89	departed, department, departments, departments'
seminars	8	132	1.40	seminar, seminars
opportunities	13	113	1.20	opportunities, opportunity
culture	7	78	0.83	cultural, culture, culture', cultures
school	6	75	0.80	school, schools
community	9	74	0.79	communal, communicate, communicated, communication, communities, community
discuss	7	68	0.72	discuss, discussed, discusses, discussing, discussion, discussions
support	7	63	0.67	support, supported, supporting, supportive, supports
Little	6	56	0.59	little
group	5	54	0.57	group, groups
works	5	52	0.55	worked, working, 'working, works
within	6	50	0.53	within

Table 13 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research culture".

Students
<ul style="list-style-type: none"> <li>I was able to interact and work along with students from multiple domains engaged with the same project.</li> <li>Within the wing there is good collaboration, but little with other wings and graduate students</li> <li>As I'm not based in my department, opportunities to meet other research students and learn from them are few.</li> <li>toxic PhD student atmosphere across department.</li> <li>There is no other research student doing the type of research that I am doing so that is why I can't really discuss it in a meaningful way with others.</li> </ul>
Departments
<ul style="list-style-type: none"> <li>Research seems to be kept to each lab and not spread around the department.</li> <li>Department have conducted small seminars, conferences and meetings where we can discuss about research or meet other students in the department.</li> <li>I feel there needs to be more direction and collaboration at the *** towards interdisciplinary research between relevant departments.</li> <li>"I have opportunities to become involved in the wider research community, beyond my department" - if I seek to do so on my own behalf.</li> <li>There are a lot of conferences run by the department, but I run a conference of my own with a and the department doesn't provide us with any support or funding.</li> </ul>

Table 14 Sample of reference extracts from the two most frequently occurring words to the question relating to "Research culture".

### 3.4.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Research culture* from 2018 to 2019 and domicile group are explored, Table 15 and Table 16, respectively, to determine where or not trends exist in how student’s respond to the question.

Year of fieldwork	Research culture		Total
	Departments	Students	
2018	158 (44.9%)	194 (55.1%)	352 (100.0%)
2019	114 (47.5%)	126 (52.5%)	240 (100.0%)
Total	272 (45.9%)	320 (54.1%)	592 (100.0%)

Table 15 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Research culture”.

For *Research culture*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Departments* and *Students* is not statistically significant ( $p = 0.5310$ ). Therefore, from 2018 to 2019 the most frequently occurring words summarised in Table 13 - Table 14 are consistent.

Domicile group	Research culture		Total
	Departments	Students	
Irish	190 (43.7%)	245 (56.3%)	435 (100.0%)
Non-Irish	82 (52.2%)	75 (47.8%)	157 (100.0%)
Total	272 (45.9%)	320 (54.1%)	592 (100.0%)

Table 16 Matrix coding of the relationship between Domicile group and the two most frequently occurring words to the question relating to “Research culture”.

Although there is a trend that more Irish students reference *students* when responding to the questions relating to *Research Culture* and Non-Irish students reference *Departments*, based on the data to date this difference is not statistically significant ( $p = 0.0653$ ). Hence, Irish students and non-Irish students are consistent in their the two most frequently occurring words to the question relating to *Research culture*.

### 3.5 Induction, progression arrangements and assessment

#### 3.5.1 Word frequency analysis

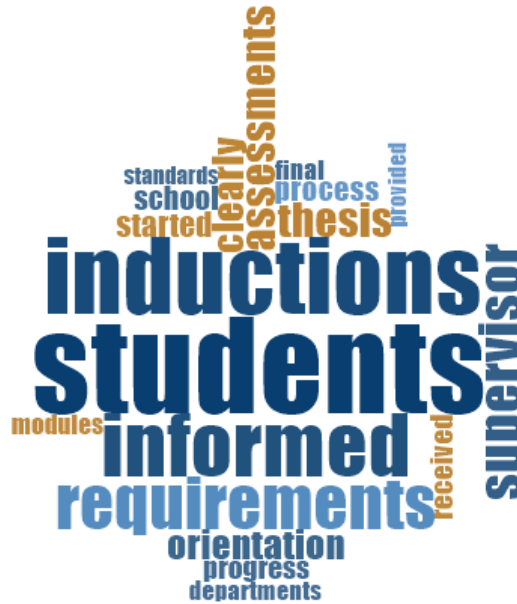


Figure 5 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Induction, progression arrangements and assessment”.

From Figure 5 and Table 17 it is apparent that the two most frequently occurring words from the question relating to *Induction, progression arrangements and assessment* are *students* and *inductions*. Table 18 provides a sample of references extracted from students highlighting predominantly a lack of induction in many instances.

Word	Length	Count	Weighted Percentage (%)	Similar Words
students	8	217	2.39	student, students, students'
inductions	10	173	1.91	induction, inductions, 'inductions'
informed	8	145	1.60	inform, informal, information, informations, informative, informed, informing, informs
requirements	12	120	1.32	require, required, requirement, requirements, requires
supervisor	10	104	1.15	supervisor, supervisors, supervisors'
assessments	11	82	0.90	assess, assessed, assessment, assessments
Thesis	6	82	0.90	thesis
clearly	7	75	0.83	clear, clearly
Orientation	11	72	0.79	orientation, orientations
Started	7	61	0.67	start, started, starting, starts
Process	7	60	0.66	process, processed, processes
School	6	58	0.64	school, schools
Received	8	56	0.62	receive, received, receiving

Progress	8	55	0.61	progress, progressing, progression, progressive
Final	5	52	0.57	final, finalized, finally
Modules	7	50	0.55	module, modules
departments	11	49	0.54	department, departments
provided	8	46	0.51	provide, provided, provides, providing
standards	9	45	0.50	standard, standards

Table 17 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Induction, progression arrangements and assessment".

Students
<ul style="list-style-type: none"> <li>• There was no formal induction, and many of the required steps I needed to take, I had to ask fellow students about.</li> <li>• Induction process needs to improve so that students understand who to go to with what queries.</li> <li>• Defining the procedure of Ph.D. at the start of the program would be great. e.g I did not have any idea about DSP or transfer assessment.</li> <li>• As far as I am aware there is no induction for PhD students.</li> <li>• There was one orientation for PhD students. Unfortunately the registration process was so long we had to do it</li> </ul>
Inductions
<ul style="list-style-type: none"> <li>• The induction has helped me to go through this postgraduate journey.</li> <li>• Did not really have an orientation or induction upon my arrival here for the PhD.</li> <li>• I think that this area is very vague. The information is not easily accessible. If you need information you must email the research dept (the admin staff there are brilliant) but there was no induction or induction pack.</li> <li>• Very poor induction and communication throughout my Phd from my school and supervisor.</li> <li>• Never received any sort of induction/orientation nor any guidance to *** wider resources.</li> </ul>

Table 18 Sample of reference extracts from the two most frequently occurring words to the question relating to "Induction, progression arrangements and assessment".

### 3.5.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Induction, progression arrangements and assessment* from 2018 to 2019 and mode of study are explored, Table 19 and Table 20, respectively, to determine where or not trends exist in how student’s respond to the question.

Year of fieldwork	Induction, progression arrangements and assessment		Total
	Inductions	Students	
2018	98 (45.6%)	117 (54.4%)	215 (100.0%)
2019	75 (42.9%)	100 (57.1%)	175 (100.0%)
Total	173 (44.4%)	217 (55.6%)	390 (100.0%)

Table 19 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Induction, progression arrangements and assessment”.

For *Induction, progression arrangements and assessment*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Inductions* and *Students* is not statistically significant ( $p = 0.5902$ ) – i.e., there is no change in student’s most prevalent word use from 2018 to 2019 in response to the question relating to *Induction, progression arrangements and assessment*.

Mode of study	Induction, progression arrangements and assessment		Total
	Inductions	Students	
Full-time	155 (43.5%)	201 (56.5%)	356 (100.0%)
Part-time/Remote	18 (52.9%)	16 (47.1%)	34 (100.0%)
Total	173 (44.4%)	217 (55.6%)	390 (100.0%)

Table 20 Matrix coding of the relationship between Mode of study and the two most frequently occurring words to the question relating to “Induction, progression arrangements and assessment”.

For *Induction, progression arrangements and assessment* and the two most frequently occurring words – *Inductions* and *Students* – the difference in distribution of references between mode of study types is not statistically significant ( $p = 0.2918$ ) – i.e., full-time students do not differ significantly in their two most frequently occurring words to the question relating to *Induction, progression arrangements and assessment* to part-time/remote students.

### 3.6 Development opportunities (including teaching / demonstrating)

#### 3.6.1 Word frequency analysis



Figure 6 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Development opportunities (including teaching / demonstrating)”.

From Figure 6 and Table 21 it is apparent that the two most frequently occurring words from the question relating to *Development opportunities (including teaching / demonstrating)* are *teaching* and *students*. Table 22 provides a sample of references extracted from students highlighting a mix of teaching being a great experience to there being no opportunity to avail of same.

Word	Length	Count	Weighted Percentage (%)	Similar Words
teaching	8	661	4.44	teach, teaches, teaching, teachings
students	8	383	2.58	student, student', students, students'
demonstrating	13	317	2.13	demonstrate, demonstrated, demonstrating, demonstration, demonstrations, demonstrator, demonstrators, demonstrators'
opportunities	13	200	1.34	opportunities, opportunity, opportunity'
Hours	5	156	1.05	hourly, hours, hours'
trainings	9	156	1.05	train, trained, training, trainings
modules	7	126	0.85	module, modules
experiments	11	124	0.83	experience, experiences, experiment, experiments
departments	11	115	0.77	department, departments
availed	7	110	0.74	avail, availability, available, available', availed, 'availed, availing
support	7	106	0.71	support, supported, supporting, supportive, supports
given	5	100	0.67	given
lecturing	9	99	0.67	lecture, lectured, lecturer, lecturers, lectures, lecturing

development	11	78	0.52	develop, develope, developed, developement, developing, development
university	10	76	0.51	universal, universities, university
requires	8	75	0.50	require, required, requirement, requirements, requires, requiring

Table 21 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Development opportunities (including teaching / demonstrating)".

Teaching
<ul style="list-style-type: none"> <li>• Found teaching to be a bit of a distraction from my own research but I am glad for the opportunity and experience of teaching in an university environment.</li> <li>• I would like to have more opportunities for teaching / demonstrating. I have done it once or twice because I personally asked. As a Post graduate student I feel I have more to give and more opportunities should be created for us to teach.</li> <li>• Further training on classroom management and development of curriculum would be useful.</li> <li>• I'd be happy to get involved more in teaching - there are no many oppourtunities for my qualifications in my department/institution.</li> <li>• Teaching or demonstrating did not happened because the supervisor will not allow it.</li> </ul>
Students
<ul style="list-style-type: none"> <li>• There is no formal tutoring training, which is fine for me because I have learned it in previous institutions/jobs. But I worry about the demonstration quality overall for students.</li> <li>• For part time students it would be great to have some workshops after 5pm.</li> <li>• Other research students have been accorded more teaching than myself in the last two years which seems very unfair.</li> <li>• Nothing for part time PhD students in school of education</li> <li>• I am an international student paying 15000 euros/year. I should not waste time demonstrating for anybody.</li> </ul>

Table 22 Sample of reference extracts from the two most frequently occurring words to the question relating to "Development opportunities (including teaching / demonstrating)".

### 3.6.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Development opportunities (including teaching / demonstrating)* from 2018 to 2019 and programme type are explored, Table 23 and Table 24, respectively, to determine where or not trends exist in how student’s respond to the question.

Year of fieldwork	Development opportunities (including teaching / demonstrating)		Total
	Teaching	Students	
2018	366 (64.3%)	203 (35.7%)	569 (100.0%)
2019	295 (62.1%)	180 (37.9%)	475 (100.0%)
Total	661 (63.3%)	383 (36.7%)	1044 (100.0%)

Table 23 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Development opportunities (including teaching / demonstrating)”.

For *Development opportunities (including teaching / demonstrating)*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Teaching* and *Students* is not statistically significant ( $p = 0.4590$ ). From Table 23 it is apparent that student’s two most frequently occurring words to the question relating to development opportunities focuses mainly on *Teaching* in both 2018 and 2019.

Programme type	Development opportunities (including teaching / demonstrating)		Total
	Teaching	Students	
Masters	74 (64.3%)	41 (35.7%)	115 (100.0%)
PhD	587 (63.2%)	342 (36.8%)	929 (100.0%)
Total	661 (63.3%)	383 (36.7%)	1044 (100.0%)

Table 24 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to “Development opportunities (including teaching / demonstrating)”.

For *Development opportunities (including teaching / demonstrating)* and the two most frequently occurring words – *Teaching* and *Students* – the difference in distribution of references between programme types is not statistically significant ( $p = 0.8074$ ). Hence, both Masters and PhD students are aligned in the two most frequently occurring words to the question relating to *Development opportunities (including teaching / demonstrating)*.

### 3.7 Research skills development

#### 3.7.1 Word frequency analysis



Figure 7 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research skills development”.

From Figure 7 and Table 25 it is apparent that the two most frequently occurring words from the question relating to *Research skills development* are *development/skills* and *training*. Table 26 provides a sample of references extracted from students highlighting a lack of development and training available to supervisors and training being excellent in relation to this development.

Word	Length	Count	Weighted Percentage (%)	Similar Words
development	11	80	2.46	develop, developed, developing, development
skills	6	80	2.46	skill, skills
training	8	34	1.04	training, trainings
integrity	9	33	1.01	integrity, integrity'
students	8	33	1.01	student, students
modules	7	32	0.98	module, modules
learning	8	30	0.92	learn, learned, learning, learnings
course	6	29	0.89	course, courses
supervisor	10	29	0.89	supervisor, supervisors
ethics	6	28	0.86	ethical, ethically, ethics
institution	11	26	0.80	institute, institution, institutions
understanding	13	25	0.77	understand, understanding, understands
improved	8	22	0.68	improve, improved, improvement, improvements, improving

university	10	22	0.68	universities, university
programme	9	19	0.58	programme, programmes
degree	6	17	0.52	degree

Table 25 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research skills development”.

Development/skills*
<ul style="list-style-type: none"> <li>The research skills development is a longlife and continuous learning.</li> <li>I feel I am still developing these skills.</li> <li>None of this skills were developed thanks to my supervisor or my research team.</li> <li>Supervisors have been great help in developing such skills.</li> <li>These research skills development came from my activities as an individual researcher, and does not reflect on ***.</li> </ul>
Training
<ul style="list-style-type: none"> <li>The research skills training programme in my institution has been excellent in providing workshops and training.</li> <li>Training courses and supervision have been excellent at engaging these points.</li> <li>There has been no training offered.</li> <li>Research methodology has to be strengthened, therefore it needs more trainings on this.</li> <li>Lack of training in qualitative research methods.</li> </ul>

Table 26 Sample of reference extracts from the two most frequently occurring words to the question relating to “Research skills development”.

\* Referenced equally with same weighted percentage, see Table 25.

### 3.7.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Research skills development* from 2018 to 2019 and programme type are explored, Table 27 and Table 28, respectively, to determine where or not trends exist in how student’s respond to the question.

Year of fieldwork	Research skills development		Total
	Development/skills	Training	
2018	42 (77.8%)	12 (22.2%)	54 (100.0%)
2019	38 (63.3%)	22 (36.7%)	60 (100.0%)
Total	80 (70.2%)	34 (29.8%)	114 (100.0%)

Table 27 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Research skills development”.

For *Research skills development*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Development/skills* and *Training* is not statistically significant ( $p = 0.0923$ ). Hence, student’s two most frequently occurring words to the question relating to *Research skills development* are consistent from 2018 to 2019.

Programme type	Research skills development		Total
	Development/skills	Training	
Masters	8 (80.0%)	2 (20.0%)	10 (100.0%)
PhD	72 (69.2%)	32 (30.8%)	104 (100.0%)
Total	80 (70.2%)	34 (29.8%)	114 (100.0%)

Table 28 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to “Research skills development”.

For *Research skills development* and the two most frequently occurring words – *Development/skills* and *Training* – the difference in distribution of references between programme types is not statistically significant ( $p = 0.4771$ ). Hence, both Masters and PhD students are aligned in their two most frequently occurring words to the question relating to *Research skills development*.

### 3.8 Development of other transferable skills

#### 3.8.1 Word frequency analysis



Figure 8 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Development of other transferable skills”.

From Figure 8 and Table 29 it is apparent that the two most frequently occurring words from the question relating to *Development of other transferable skills* are *development* and *skills*. Table 30 provides a sample of references extracted from students that vary from a supervisor being a great support in developing skills, to there being no opportunity and students being left to acquire independently.

Word	Length	Count	Weighted Percentage (%)	Similar Words
development	11	86	3.60	develop, developed, developing, development, developments
skills	6	78	3.26	skill, skilled, skills
managing	8	44	1.84	manage, 'manage, managed, management, manager, manages, managing
project	7	43	1.80	project, project', projects
professional	12	39	1.63	professional, professionally
communication	13	35	1.46	communicate, communicated, communicating, communication, communications, community
supervisors	11	29	1.21	supervisor, supervisors, supervisors'
students	8	28	1.17	student, students
transferable	12	27	1.13	transfer, transferable
opportunities	13	24	1.00	opportunities, opportunity
networking	10	22	0.92	network, networking, networks
career	6	19	0.79	career, careers
programme	9	17	0.71	programme
Start	5	16	0.67	start, started, starting
support	7	16	0.67	support, supported, supports
experiences	11	14	0.59	experience, experiences
Years	5	14	0.59	years
modules	7	13	0.54	module, modules
audience	8	12	0.50	audience, audiences
university	10	12	0.50	universities, university

Table 29 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Development of other transferable skills".

Development
<ul style="list-style-type: none"> <li>• I have been given many opportunities by my advisory supervisor to shadow her in her work and to assist when possible. This has really been a huge opportunity for my professional development in above areas.</li> <li>• I am still developing these skills, as I am only beginning year 2, but I have many opportunities to do so.</li> <li>• All developed before the PhD (previous research, practical experience) All improving with the PhD.</li> <li>• I arrived at the programme fairly competent to start with. There has been no mentoring with regard to career development, or networking. I've just continued to do it myself.</li> <li>• I have had to be very self driven in this regard, my supervisor is averse to my 'wasting time' with courses or developing non essential skills.</li> </ul>
Skills
<ul style="list-style-type: none"> <li>• As an older student, my professional and communication skills have long been a part of my career.</li> <li>• These skills are necessary to manage your pathway towards your research programme.</li> <li>• Development of transferable skills, especially project planning, has been entirely left up to myself, with my supervisors having very little input.</li> <li>• All self taught or taught through peers. No university programme or tools to facilitate these skills.</li> <li>• I was not offered much guidance in regards to developing transferable skills, I'm still unclear as to what my transferable skills are.</li> </ul>

Table 30 Sample of reference extracts from the two most frequently occurring words to the question relating to “Development of other transferable skills”.

### 3.8.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Development of other transferable skills* from 2018 to 2019 and domicile group are explored, Table 31 and Table 32, respectively, to determine where or not trends exist in how student’s respond to the question.

Year of fieldwork	Development of other transferable skills		Total
	Development	Skills	
2018	40 (46.5%)	46 (53.5%)	86 (100.0%)
2019	46 (59.0%)	32 (41.0%)	78 (100.0%)
Total	86 (52.4%)	78 (47.6%)	164 (100.0%)

Table 31 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Development of other transferable skills”.

For *Development of other transferable skills*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Development* and *Skills* is not statistically significant ( $p = 0.1105$ ). Hence, the two most frequently occurring words displayed in the word frequency analysis are consistent from 2018 to 2019.

Domicile group	Development of other transferable skills		Total
	Development	Skills	
Irish	62 (51.2%)	59 (48.8%)	121 (100.0%)
Non-Irish	24 (55.8%)	19 (44.2%)	43 (100.0%)
Total	86 (52.4%)	78 (47.6%)	164 (100.0%)

Table 32 Matrix coding of the relationship between Domicile group and the two most frequently occurring words to the question relating to “Development of other transferable skills”.

For *Development of other transferable skills* and the two most frequently occurring words – *Development* and *Skills* – the difference in distribution of references between domicile group is not statistically significant ( $p = 0.6059$ ) – i.e., the results from the two most frequently occurring words are consistent between Irish and non-Irish students.

### 3.9 Student / staff responsibilities and supports

#### 3.9.1 Word frequency analysis



Figure 9 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Student / staff responsibilities and supports”.

From Figure 9 and Table 33 it is apparent that the two most frequently occurring words from the question relating to *Student / staff responsibilities and supports* are *student*’ and *support*. Table 34 provides a sample of references extracted from students ranging from a positive experience with regards supports to none being existent. Some students state that although supports are available, they are not overly visible and apparent.

Word	Length	Count	Weighted Percentage (%)	Similar Words
students'	9	305	4.25	student, 'student', students, students'
support	7	163	2.27	support, supported, supporting, supportive, supports
feedback	8	82	1.14	feedback
staff	5	75	1.04	staff, 'staff'
supervisor	10	73	1.02	supervisor, supervisors, supervisors'
departments	11	61	0.85	department, departments
responsive	10	53	0.74	response, responses, responsibilities, responsibility, responsible, responsive, responsiveness
issues	6	51	0.71	issue, issues
institution	11	45	0.63	institute, institution, institutional, institutions
school	6	42	0.58	school
university	10	41	0.57	universities, university
available	9	39	0.54	avail, available, availed
postgraduate	12	37	0.52	postgraduate, postgraduates

Table 33 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Student / staff responsibilities and supports”.

Students'
<ul style="list-style-type: none"> <li>• It would be better if PhD students are allowed to parking on campus.</li> <li>• Supports are not available/suited to part time postgraduate students.</li> <li>• I find *** to be a very supportive environment. I include library staff and other students.</li> <li>• *** neglects its postgraduate students and its duties towards them. It does not provide them with proper accommodation supports during this housing crisis or proper student spaces.</li> <li>• There need to be equal supports available to the research students as there are undergrads.</li> </ul>
Support
<ul style="list-style-type: none"> <li>• There are are college supports. there are no supports within the department.</li> <li>• The support from the lecturing staff has been excellent.</li> <li>• Student supports should be made more visible. I found the supports that I have needed through searching, but it was not obvious what supports are available outside of research support in the library and use of the arena.</li> <li>• Not very aware of what is available in terms of Recreation and similar supports in ***.</li> <li>• Mental health supports are virtually non existent.</li> </ul>

Table 34 Sample of reference extracts from the two most frequently occurring words to the question relating to “Student / staff responsibilities and supports”.

### 3.9.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Student / staff responsibilities and supports* from 2018 to 2019 and mod of study are explored, Table 35 and Table 36, respectively, to determine where or not trends exist in how student's respond to the question.

Year of fieldwork	Student / staff responsibilities and supports		Total
	Students'	Support	
2018	158 (64.5%)	87 (35.5%)	245 (100.0%)
2019	147 (65.9%)	76 (34.1%)	223 (100.0%)
Total	305 (65.2%)	163 (34.8%)	468 (100.0%)

Table 35 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Student / staff responsibilities and supports".

For *Student / staff responsibilities and supports*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Students'* and *Support* is not statistically significant ( $p = 0.7458$ ) – i.e., the most prevalent frequently occurring words for both 2018 and 2019 is *students'*.

Mode of study	Student / staff responsibilities and supports		Total
	Students'	Support	
Full-time	263 (65.1%)	141 (34.9%)	404 (100.0%)
Part-time/Remote	42 (65.6%)	22 (34.4%)	64 (100.0%)
Total	305 (65.2%)	163 (34.8%)	468 (100.0%)

Table 36 Matrix coding of the relationship between Mode of study and the two most frequently occurring words to the question relating to "Student / staff responsibilities and supports".

For *Student / staff responsibilities and supports* and the two most frequently occurring words – *Students'* and *Support* – the difference in distribution of references between mode of study is not statistically significant ( $p = 0.9346$ ). Hence, both full-time and part-time/remote students are aligned with the two most frequently occurring words to the question relating to *Student / staff responsibilities and supports*.

### 3.10 Personal outlook

#### 3.10.1 Word frequency analysis



Figure 10 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Personal outlook”.

From Figure 10 and Table 37 it is apparent that the two most frequently occurring words from the question relating to *Personal outlook* are *students* and *support*. Table 38 provides a sample of references extracted from students highlighting issues with lack of financial support and that the cost of living not mapping to student pay, with instances of where support for students is lacking in general.

Word	Length	Count	Weighted Percentage (%)	Similar Words
students	8	118	2.24	student, students, students'
support	7	55	1.04	support, supported, supportive, supports
school	6	49	0.93	school, schooling
department	10	48	0.91	department, departments
balance	7	37	0.70	balance, balancing
studies	7	37	0.70	studies, study, studying
supervisors	11	35	0.66	supervisor, supervisors
graduate	8	30	0.57	graduate, graduates, graduation
programme	9	29	0.55	programme, 'programme', programmer, programmes
institution	11	29	0.55	institute, institutes, institution, institutional
academic	8	27	0.51	academic, academics

Table 37 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Personal outlook”.

<p>Students</p> <ul style="list-style-type: none"> <li>• Rents are too high. You could consider to increase PhD students salaries. We are smart hardworking people who deserve better accommodations!</li> <li>• Financial supports are my biggest worry - its tough to be a fulltime student with no income when you live with rent and bills.</li> <li>• Honestly, I don't think that student support can help when dealing with low qualified lecturers. The institution does not humbly assess the quality of its lecturers, that in some cases is poor.</li> <li>• I am a volunteer with a national mental health charity and I would have concerns for postgraduate students struggling with things having an outlet to talk about things with their peers.</li> <li>• There should be more support for PhD students past having coffee mornings and drop in sessions. An ombudsmen perhaps.</li> </ul>
<p>Support</p> <ul style="list-style-type: none"> <li>• PhDs = pain with little meaningful support outside of academic problems from the university.</li> <li>• *** as an institution has not valued or supported my study - work life balance requests. equally, my research contribution as a part time post grad working full time has been undermined by line management in work remit.</li> <li>• There have definitely been times, among myself and others in my school, where we have questioned why we are doing a PhD as often times we are not supported at a university level and our problems concerning finances are trivialised.</li> <li>• My fellow PhD students have been a massive support to me without whom I'm not sure I'd still be in the programme. I would worry for students in smaller research groups that don't have that support.</li> <li>• Overall, my experience at *** has been positive but there are issues with pay equality, research ambience and support programmes/initiatives from my department.</li> </ul>

*Table 38 Sample of reference extracts from the two most frequently occurring words to the question relating to “Personal outlook”.*

### 3.10.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Personal outlook* from 2018 to 2019 and programme type are explored, Table 39 and Table 40, respectively, to determine where or not trends exist in how student’s respond to the question.

Year of fieldwork	Personal outlook		Total
	Students	Support	
2018	0 (0.0%)	0 (0.0%)	0 (0.0%)
2019	118 (68.2%)	55 (31.8%)	173 (100.0%)
Total	118 (68.2%)	55 (31.8%)	173 (100.0%)

Table 39 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Personal outlook”.

For *Personal outlook*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Students* and *Support* could not be determined due to there being no references to the stemmed variants in 2018.

Programme type	Personal outlook		Total
	Students	Support	
Masters	20 (74.1%)	7 (25.9%)	27 (100.0%)
PhD	98 (67.1%)	48 (32.9%)	146 (100.0%)
Total	118 (68.2%)	55 (31.8%)	173 (100.0%)

Table 40 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to “Personal outlook”.

For *Personal outlook* and the two most frequently occurring words – *Students* and *Support* – the difference in distribution of references across programme types is not statistically significant ( $p = 0.4761$ ) – i.e., the two most frequently occurring words to the question relating to *Personal outlook* are similar for both Masters and PhD students.



project	7	176	0.91	project, projects
teaching	8	174	0.90	teach, teaching, 'teaching
works	5	154	0.80	worked, working, works
modules	7	143	0.74	module, modules
available	9	142	0.74	avail, availability, available, availing
networking	10	140	0.73	network, networking, networks
supervision	11	139	0.72	supervise, supervised, supervisor, supervising, supervision
access	6	137	0.71	access, accessibility, accessible, accessing
community	9	133	0.69	communal, communicate, communicating, communication, communications, communicator, community
conferences	11	132	0.69	conference, conferences
knowledgeable	13	120	0.62	knowledgable, knowledge, knowledgeable
departments	11	117	0.61	department, departments
facilities	10	109	0.57	facilities, facility
management	10	104	0.54	manage, manageable, managed, management, managing, managment
environment	11	103	0.53	environment
field	5	102	0.53	field, fields
independently	13	102	0.53	independance, independant, independence, independent, independently
professional	12	101	0.52	professional, professionalism, professionally, professionals
presenting	10	96	0.50	present, presentation, presentations, presented, presenting

Table 41 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Aspects / elements of research degree programme are most valuable".

Supervisors'
<ul style="list-style-type: none"> <li>• Supervisor and department support.</li> <li>• Interactions with supervisors; opportunities to participate in postgrad seminars.</li> <li>• Supervisor's approachability and accountability, approachable faculty, and good department.</li> <li>• Without doubt the support, approachability, enthusiasm and feedback from my principal supervisor.</li> <li>• Independent thinking and the supervisor - student interactions.</li> </ul>
Skills
<ul style="list-style-type: none"> <li>• I have self-taught myself a number of valuable technical skills and published in a high impact journal.</li> <li>• The transferable skill and the opportunity to engage with peers and lay audiences.</li> <li>• The different analytical skills learned during the program and the opportunity to meet high level of researches in my field.</li> <li>• Structure of the programme that allows students to have time for both research and skills/knowledge development.</li> <li>• learning new skills Critical thinking.</li> </ul>

Table 42 Sample of reference extracts from the two most frequently occurring words to the question relating to "Aspects / elements of research degree programme are most valuable".

### 3.11.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Aspects / elements of research degree programme are most valuable* from 2018 to 2019 and gender are explored, Table 43 and Table 44, respectively, to determine where or not trends exist in how student's respond to the question.

Year of fieldwork	Aspects / elements of research degree programme are most valuable		Total
	Skills	Supervisors'	
2018	285 (45.0%)	349 (55.0%)	634 (100.0%)
2019	235 (44.0%)	299 (56.0%)	534 (100.0%)
Total	520 (44.5%)	648 (55.5%)	1168 (100.0%)

Table 43 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Aspects / elements of research degree programme are most valuable".

For *Aspects / elements of research degree programme are most valuable*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Skills* and *Supervisors'* is not statistically significant ( $p = 0.7461$ ). Hence, consistently for both 2018 and 2019 *Supervisors'* are stated as the valuable aspect of the research programme.

Gender	Aspects / elements of research degree programme are most valuable		Total
	Skills	Supervisors'	
Female	350 (45.0%)	427 (55.0%)	777 (100.0%)
Male	170 (43.5%)	221 (56.5%)	391 (100.0%)
Total	520 (44.5%)	648 (55.5%)	1168 (100.0%)

Table 44 Matrix coding of the relationship between Gender and the two most frequently occurring words to the question relating to "Aspects / elements of research degree programme are most valuable"

For *Aspects / elements of research degree programme are most valuable* and the two most frequently occurring words – *Skills* and *Supervisors'* – the difference in distribution of references with respect to gender is not statistically significant ( $p = 0.6111$ ). Hence, both males and females are in agreement that the supervisor is the most valuable aspect of the research programme.

### 3.12 Aspects of research degree experience could be improved

#### 3.12.1 Word frequency analysis



Figure 12 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Aspects of research degree experience could be improved".

From Figure 12 and Table 45 it is apparent that the two most frequently occurring words from the question relating to *Aspects of research degree experience could be improved* are *students* and *support*. Table 46 provides a sample of references extracted from students highlighting supports for student's in general needing improvement – e.g., support in understanding regulations to career aspects to economical, etc.

Word	Length	Count	Weighted Percentage (%)	Similar Words
students	8	819	2.77	student, students, students', 'students'
support	7	468	1.58	support, supported, supporting, supportive, supports

supervisors	11	362	1.22	supervisor, supervisors, supervisors', 'supervisors'
funds	5	329	1.11	funded, funding, fundings, funds
modules	7	285	0.96	module, modules
trainings	9	269	0.91	train, trained, training, trainings
departments	11	228	0.77	department, departments
better	6	213	0.72	better
opportunities	13	207	0.70	opportunities, opportunity
community	9	189	0.64	communal, communicate, communicated, communicating, communication, communications, communities, 'communities', community
available	9	188	0.64	avail, availabilities, availability, available
access	6	180	0.61	access, accessed, accessibility, accessible, accessing
supervision	11	172	0.58	supervise, supervises, supervising, supervision, supervisions
facilities	10	169	0.57	facilities, facility, facilities
teaching	8	164	0.55	teach, teaching
improved	8	163	0.55	improve, improved, improvement, improvements, improves, improving
skills	6	161	0.54	skill, skills, skills'
structured	10	152	0.51	structure, structured, structurent, structures, structuring

Table 45 NVivo word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Aspects of research degree experience could be improved”.

<p><b>Students</b></p> <ul style="list-style-type: none"> <li>• There needs to be a clear code of conduct for PhD students and supervisors - If students or supervisors do not comply with the code of conduct issues must be resolved in a prompt manner - Accountability is seriously lacking in my University. - Foster more supportive research environments. Clicky groups, exclusion and jealousy do not create a professional research environment.</li> <li>• Greater opportunities to get to know other research students.</li> <li>• Career aspects in relation to mature student.</li> <li>• Statistical analysis help/statistician for postgraduate students Knowledge of equipment here and training that could be completed Office infrastructure.</li> <li>• Information about regulations for PhD students could be communicated in a better way.</li> </ul> <p><b>Support</b></p>
--

- Healthcare support services for students.
- Post-submission support by College, and career guidance specific to area (i.e. Humanities, not STEM).
- Every. single. part. There is no support. Supervision is terrible.
- Economical support to attend conferences.
- Support for part time students would be helpful.
- Support from supervisors.

Table 46 Sample of reference extracts from the two most frequently occurring words to the question relating to “Aspects of research degree experience could be improved”.

### 3.12.2 Matrix coding and statistical inference

In this subsection, the distribution of the two most frequently occurring words in response to the question relating to *Aspects of research degree experience could be improved* from 2018 to 2019 and HEI type are explored, Table 47 and Table 48, respectively, to determine where or not trends exist in how student’s respond to the question.

Year of fieldwork	Aspects / elements of research degree programme could be improved		Total
	Students	Support	
2018	431 (61.4%)	271 (38.6%)	702 (100.0%)
2019	388 (66.3%)	197 (33.7%)	585 (100.0%)
Total	819 (63.6%)	468 (36.4%)	1287 (100.0%)

Table 47 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Aspects / elements of research degree experience could be improved”.

For *Aspects of research degree experience could be improved*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Students* and *Support* is not statistically significant ( $p = 0.0672$ ). Hence, for both years of data collection student support is prevalent as the aspect of the research experience that needs improving.

HEI type	Aspects of research degree experience could be improved		Total
	Students	Support	
Institutes of Technology	150 (65.2%)	80 (34.8%)	230 (100.0%)
Universities	648 (63.3%)	375 (36.7%)	1023 (100.0%)
Other Institutions	21 (61.8%)	13 (38.2%)	34 (100.0%)
Total	819 (63.6%)	468 (36.4%)	1287 (100.0%)

Table 48 Matrix coding of the relationship between HEI type and the two most frequently occurring words to the question relating to “Aspects of research degree experience could be improved”.

For *Aspects of research degree experience could be improved* and the two most frequently occurring words – *Students* and *Support* – the difference in distribution of references with respect to HEI type is not statistically significant ( $p = 0.8445$ ). Hence, student support being an aspect of the research experience that could be improved is consistent across HEI types.

## 4. Summary and Conclusion

In this project, a roadmap is presented to provide a mechanism for qualitative data analysis on PGR StudentSurvey.ie. The roadmap has six steps:

1. Data cleaning;
2. Data import;
3. Word frequency analysis;
4. Text coding and Reference extracting;
5. Matrix coding;
6. Statistical inference.

The six steps to the roadmap are applied to the 11 open-ended question responses of the PGR StudentSurvey.ie data from 2018 and 2019 using NVivo and RStudio. As summarised in Table 4 the results across packages are consistent with the exception of the responses to *Research infrastructure and facilities* due to how each package manages stemmed variants of a word. Which package to use will be very much dependent on the user with a sample of advantages and disadvantages to each package outlined in Table 1. NVivo has a lot of additional user-friendly functionality than what is demonstrated in this project – e.g., hierarchy charts, mind maps, comparison diagrams, etc. Although, most of the functionality in NVivo can be replicated in RStudio, determining the appropriate script in the latter is a challenge. But this is potentially outweighed by RStudio being open-source as opposed to NVivo having a cost implication.

The overall structure to REVEAL-PGR is based on the five steps outlined Feng & Horenstein (2019), however, it is the author's belief that the sixth step – *Statistical inference* – offers a synergy between qualitative analysis with quantitative analysis with results that should complement the narrative around the qualitative analysis. The statistical inference step enables the possibility to determine if trends exist in students' responses for a particular year and/or demographic – Higher Education Institution (HEI) type, Gender, Domicile group, Mode of study and Programme type.

There are two main objectives to this project:

1. Qualitative analysis on the 11 open-ended questions from 2018 and 2019 data;
2. Introduce a roadmap to promote the accessibility of the analysis of the unstructured qualitative data emerging from PGR StudentSurvey.ie.

In relation to the latter, irrespective of package used, there is a lot of potential to expand the capability of the roadmap. In this project, six sources of demographic data are used – i.e., Year of fieldwork, HEI type, Gender, Domicile group, Mode of study, and Programme type - however there are numerous closed-ended questions from PGR StudentSurvey.ie that association/trends between the references from the qualitative analysis could be explored. In this project the references extracted, matrix coding and statistical inference are based on the two most frequently occurring words in response to a question. However, there are many more frequently used words (with a weighed percentage of at least 0.5%) that could be explored for other potentially interesting relationships. Furthermore, with

REVEAL-PGR's sixth step the options for additional rigorous statistical analysis are wide and varied – e.g., loglinear analysis, logistic regression, etc.

In relation to the qualitative analysis on the 11 open-ended questions from 2018 and 2019, it is interesting to note that references extracted are consistent and with little change from 2018 to 2019. It will be interesting to see if this stays the same in data collection 2021 given COVID-19. Across the 11 open-ended questions the most frequently occurring words are:

1. *Research infrastructure and facilities* – The two most frequently occurring words were *access* and *funding*, with the former being more prevalent in Institutes of Technology and latter in Universities and Other Institutions;
2. *Supervision* – Responses varied from supervisors being excellent to supervisors requiring training in supervision;
3. *Research culture* – Suggested a lack of collaboration between departments and some students;
4. *Induction, progression arrangements and assessment* – Students predominantly responded with there being a lack of induction training;
5. *Development opportunities (including teaching / demonstrating)* – Responses to this question were mixed between teaching being a great experience to there being no opportunity to avail of teaching/tutoring;
6. *Research skills development* – Students' response to this question were varied from a lack of development and training available to training being excellent;
7. *Development of other transferable skills* – Responses varied from supervisor being a great support in developing skills, to there being no opportunity and students being left to acquire skills independently;
8. *Student / staff responsibilities and supports* – Responses ranged from a positive experience with regards supports to none being existent;
9. *Personal outlook* – A lack of support to students' was the most prevalent response from students to this question;
10. *Aspects / elements of research degree programme are most valuable* – Supervisors were the most prevalent valuable aspect from the analysis on this question's responses;
11. *Aspects of research degree experience could be improved* – The need for additional student support was most prevalent when student's replied to this question.

When reviewing responses collectively across all 11 open-ended questions it would be appear that supervisors are the main positive influence on a student's research career, while additional student support would appear to be the most prevalent aspect that needs attention.

Overall, this project showcases the accessibility of a roadmap to national PGR StudentSurvey.ie data through analysis of the 11 open-ended question response from 2018 and 2019. However, it will potentially be at respective HEI level, due to the diverse nature to HEI's in Ireland, where the impact and usefulness of **REVEAL – PGR (Roadmap to StudEntSurVey.ie Qualitative Data AnaLysis – PostGRaduate)** may be realised.

## 5. Acknowledgements

I would like to acknowledge Dr Siobhán Nic Fhlannchadha, StudentSurvey.ie Project Manager, for providing the necessary information regarding the project and also for her continuous support in completing the project. I would also like to express my gratitude towards Dr David Goulding, Head of Department of Mathematics, Munster Technological University – Cork, for his encouragement and allowing me the time to complete this project.

## 6. References

- Carley, K. (1993). [\*Coding choices for textual analysis: A comparison of content analysis and map analysis\*](#). Sociological Methodology, 23, 75-126.
- Estrada, S. (2017) [\*Qualitative Analysis Using R: A Free Analytic Tool\*](#). The Qualitative Project, 22(4), 956-968.
- Feng, X., & Behar-Horenstein, L. (2019) [\*Maximizing NVivo Utilities to Analyze Open-Ended Responses\*](#). The Qualitative Project 24(3), 563-571.
- Jackson, K. M., & Trochim, W. M. (2002). [\*Concept mapping as an alternative approach for the analysis of open-ended survey responses\*](#). Organizational Research Methods, 5(4), 307-336.
- QSR International <https://www.qsrinternational.com/>.
- RStudio <https://rstudio.com/>.

## Appendix RStudio output

### A.1 Research infrastructure and facilities

#### A.1.1 Word frequency analysis



Figure 13 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research infrastructure and facilities”.

Word	Length	Count	Weighted Percentage (%)
access	6	253	2.12
space	5	250	2.09
librari	7	212	1.77
facil	5	194	1.62
avail	5	141	1.18
comput	6	121	1.01
cover	5	118	0.99
univers	7	117	0.98
equip	5	115	0.96
offic	5	106	0.89
stipend	7	105	0.88
resourc	7	87	0.73
provid	6	82	0.69
confer	6	79	0.66
support	7	79	0.66
scholarship	11	77	0.64
travel	6	71	0.59
journal	7	69	0.58

studi	5	69	0.58
limit	5	69	0.58
build	5	68	0.57
depart	6	68	0.57
postgradu	9	66	0.55
campus	6	62	0.52
enough	6	61	0.51
infrastructur	13	60	0.50

Table 49 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research infrastructure and facilities".

Overlap between references to <i>Access</i> and <i>Space</i>	
•	We need more desk space and it needs to be suitable and accessible desk space and it needs to be secure for the full term of your studies (it's fine if you're asked to give it up if you don't use it or if you contravene rules, but you shouldn't be kicked out because it's only a one-year desk)
•	There is a need for the research space to have a more suitable common area and kitchen. It would also be useful to have access to a desktop at our desks.
•	I have access to a hotdesking facility. The facility is well equipped, but not having my own desk and personal shelf space for books and necessary materials is causing difficulties.
•	Not enough reading space Equipment necessary for research not accessible, have to go to *** to use equipment
•	I have truly enjoyed my time studying here in ***. One small suggestion that would be beneficial is the accessibility of printing facilities in the post grad office space in lontas.
•	Accessing facilities is not always straightforward, for example I was not allocated a desk space away from my lab and had to go through a lengthy process to receive one.

Table 50 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Research infrastructure and facilities".

### A.1.2 Matrix coding and statistical inference

Year of fieldwork	Research infrastructure and facilities		Total
	Access	Space	
2018	112 (54.6%)	93 (45.4%)	205 (100.0%)
2019	90 (49.5%)	92 (50.5%)	182 (100.0%)
Total	202 (52.2%)	185 (47.8%)	387 (100.0%)

Table 51 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Research infrastructure and facilities".

For *Research infrastructure and facilities*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Access* and *Space* is not statistically significant ( $p = 0.3592$ ).

HEI type	Research infrastructure and facilities		Total
	Access	Space	
Institutes of Technology	49 (59.0%)	34 (41.0%)	83 (100.0%)
Universities	147 (50.3%)	145 (49.7%)	292 (100.0%)
Other Institutions	6 (50.0%)	6 (50.0%)	12 (100.0%)
Total	202 (52.2%)	185 (47.8%)	387 (100.0%)

Table 52 Matrix coding of the relationship between HEI type and the two most frequently occurring words to the question relating to "Research infrastructure and facilities".

From the two most frequently occurring words, the majority of references from Institutes of Technology were related to *Access*, while the references from Universities and Other Institutions appear balanced between *Access* and *Space*. However, based on the data gathered this association is not statistically significant ( $p = 0.3713$ ).

## A.2 Supervision

### A.2.1 Word frequency analysis



Figure 14 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Supervision".

Word	Length	Count	Weighted Percentage (%)
supervisor	10	845	11.55
student	7	213	2.91
supervis	8	183	2.5
support	7	165	2.25
excel	5	80	1.09
train	5	75	1.02
provid	6	67	0.92

feedback	8	65	0.89
contact	7	58	0.79
project	7	55	0.75
experi	6	50	0.68
think	5	44	0.6
great	5	42	0.57
differ	6	41	0.56
littl	5	38	0.52
difficult	9	38	0.52
howev	5	38	0.52
direct	6	38	0.52

Table 53 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Supervision".

Overlap between references to <i>Supervisor</i> and <i>Student</i>
<ul style="list-style-type: none"> <li>• My supervisors have acted as mentors and offer regular challenges to my work. They have also argued my case when there were questions about whether to continue my study or not. This has been invaluable for me to keep going and to enter this programme as a part-time mature student</li> <li>• The supervisory supports and relationship was positive but more training or guides for research students on 'how to make the most of the supervision' would help.</li> <li>• My supervisors have very little or none input on my training courses. I actually feel that my supervisors treat me a little different from their Irish students in a negative way.</li> <li>• I feel like sometimes, supervisors are so loaded with teaching hours than they cannot provide as much attention to research students and PhDs as they would like to</li> <li>• My supervisor is excellent. However I have noticed that students whose supervisors are away on maternity leave get little to no support.</li> </ul>

Table 54 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Supervision".

### A.2.2 Matrix coding and statistical inference

Year of fieldwork	Supervision		Total
	Student	Supervisor	
2018	82 (20.1%)	326 (79.9%)	408 (100.0%)
2019	53 (17.0%)	259 (83.0%)	312 (100.0%)
Total	135 (18.8%)	585 (81.2%)	720 (100.0%)

Table 55 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Supervision".

For *Supervision*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Student* and *Supervisor* is not statistically significant ( $p = 0.3353$ ).

Gender	Supervision		Total
	Student	Supervisor	
Female	93 (19.2%)	392 (80.8%)	485 (100.0%)
Male	42 (17.9%)	193 (82.1%)	235 (100.0%)
Total	135 (18.8%)	585 (81.2%)	720 (100.0%)

Table 56 Matrix coding of the relationship between Gender and the two most frequently occurring words to the question relating to "Supervision".

For *Supervision* and the two most frequently occurring words – *Student* and *Supervisor* – the difference in distribution of references between males and females is not statistically significant ( $p = 0.7504$ ).

### A.3 Research culture

#### A.3.2 Word frequency analysis

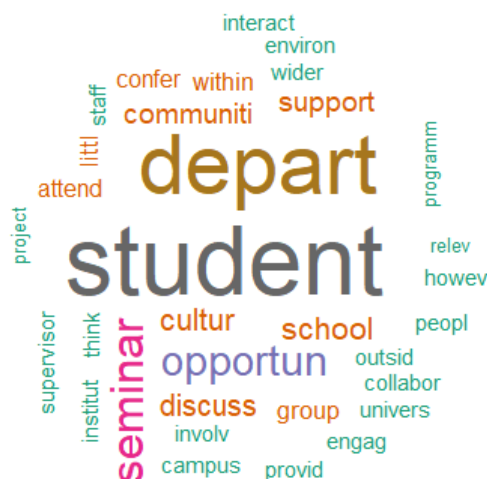


Figure 15 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research culture".

Word	Length	Count	Weighted Percentage (%)
student	7	317	4.86
depart	6	270	4.14
seminar	7	131	2.01
opportun	8	113	1.73
cultur	6	77	1.18
school	6	75	1.15
discuss	7	68	1.04
support	7	63	0.97
communiti	9	59	0.90

littl	5	56	0.86
group	5	52	0.80
within	6	50	0.77
confer	6	44	0.67
attend	6	42	0.64
think	5	39	0.60
supervisor	10	39	0.60
campus	6	39	0.60
involv	6	39	0.60
provid	6	38	0.58
outsid	6	38	0.58
engag	5	37	0.57
collabor	8	37	0.57
institut	8	37	0.57
wider	5	36	0.55
univers	7	35	0.54
peopl	5	35	0.54
environ	7	34	0.52
staff	5	34	0.52
interact	8	34	0.52
howev	5	34	0.52
project	7	33	0.51
relev	5	33	0.51
programm	8	33	0.51

Table 57 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research culture".

**Overlap between references to *Student* and *Depart(ment)***

- The department is supportive and engages students in a lot of activities in every aspect.
- Students are kept separate from the rest of the research department and are discouraged from engaging with members of staff - we have lunch separately too! This creates a tense, uncomfortable atmosphere between postgraduate students and staff.
- My department offers no student research assistance or guidance
- I think the research culture in the college is quite good. I'd like to engage more with other research students outside my department but it's difficult to find opportunities to do so, partly because there aren't many and partly because of competing time commitments.
- Within the department, more research focus between doctoral students can be made. Why don't doctoral students ever sit around a table and discuss with permanent staff?

Table 58 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Research culture".

### A.3.2 Matrix coding and statistical inference

Year of fieldwork	Research culture		Total
	Depart(ment)	Student	
2018	115 (47.9%)	125 (52.1%)	240 (100.0%)
2019	84 (48.8%)	88 (51.2%)	172 (100.0%)
Total	199 (48.3%)	213 (51.7%)	412 (100.0%)

Table 59 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Research culture".

For *Research culture*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Depart(ment)* and *Student* is not statistically significant ( $p = 0.9327$ ).

Domicile group	Research culture		Total
	Depart(ment)	Student	
Irish	139 (46.8%)	158 (53.2%)	297 (100.0%)
Non-Irish	60 (52.2%)	55 (47.8%)	115 (100.0%)
Total	199 (48.3%)	213 (51.7%)	412 (100.0%)

Table 60 Matrix coding of the relationship between Domicile group and the two most frequently occurring words to the question relating to "Research culture".

Although there is a trend that more Irish students reference *Students* when responding to the questions relating to *Research Culture* and Non-Irish students reference *Depart(ment)*, based on the data to date this difference is not statistically significant ( $p = 0.3848$ ).



programm	8	44	0.71
procedur	8	41	0.66
expect	6	40	0.65
complet	7	40	0.65
understand	10	40	0.65
degre	5	38	0.61
regard	6	35	0.56
howev	5	35	0.56
avail	5	35	0.56
littl	5	34	0.55
structur	8	33	0.53
transfer	8	33	0.53
first	5	32	0.52
given	5	32	0.52

Table 61 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Induction, progression arrangements and assessment".

#### Overlap between references to *Student* and *Induct(ion)*

- There is no formal induction program for Postgraduate students who do not start at the beginning of the semester.
- The way induction and assessment is communicated to me was through my research group. Information for this online is poorly explained through the website and orientation in particular registration for modules for Postgraduate students requires a serious makeover! The whole process is completely backwards.
- My primary supervisor is affiliated with my university, but not a full-time academic. My second supervisor is a full-time academic, but I am their first PhD student. My university does not provide induction or handbooks to help students understand requirements. As such, I am left at a loss as to what my requirements are.
- There is no induction to the PhD program, you simply walk in and start work. Whereas there should be something to help transition students and help guide them towards getting everything (ie safety, forms, etc) done. Additionally, the process of the transfer thesis hand-in is not clear. Neither is the amount of ECTS students are required to take. Most PIs do not know, let alone the students.
- Induction was run by the faculty, and was very helpful. For international or new students however, there was no induction until several months in. By then students had found their way around without help.

Table 62 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Induction, progression arrangements and assessment".

#### A.4.2 Matrix coding and statistical inference

Year of fieldwork	Induction, progression arrangements and assessment		Total
	Induct(ion)	Student	
2018	75 (48.7%)	79 (51.3%)	154 (100.0%)
2019	61 (47.3%)	68 (52.7%)	129 (100.0%)
Total	136 (48.1%)	147 (51.9%)	283 (100.0%)

Table 63 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Induction, progression arrangements and assessment".

For *Induction, progression arrangements and assessment*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Induct(ion)* and *Student* is not statistically significant ( $p = 0.9063$ ).

Mode of study	Induction, progression arrangements and assessment		Total
	Induct(ion)	Student	
Full-Time	120 (47.4%)	133 (52.6%)	253 (100.0%)
Part-time/Remote	16 (53.3%)	14 (46.7%)	30 (100.0%)
Total	136 (48.1%)	147 (51.9%)	283 (100.0%)

Table 64 Matrix coding of the relationship between Mode of study and the two most frequently occurring words to the question relating to "Induction, progression arrangements and assessment".

For *Induction, progression arrangements and assessment* and the two most frequently occurring words – *Induct(ion)* and *Student* – the difference in distribution of references between mode of study types is not statistically significant ( $p = 0.6755$ ).

## A.5 Development opportunities (including teaching / demonstrating)

### A.5.1 Word frequency analysis

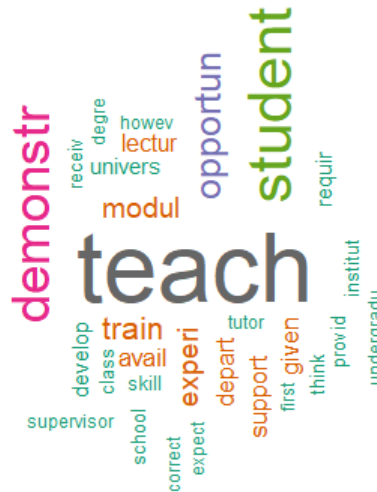


Figure 17 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Development opportunities (including teaching / demonstrating)”.

Word	Length	Count	Weighted Percentage (%)
Teach	5	660	6.57
Student	7	383	3.81
demonstr	8	317	3.16
opportun	8	200	1.99
Train	5	155	1.54
Modul	5	126	1.25
Experi	6	124	1.23
Depart	6	114	1.13
avail	5	110	1.09
support	7	106	1.06
given	5	100	1.00
lectur	6	99	0.99
develop	7	78	0.78
univers	7	76	0.76
requir	6	75	0.75
skill	5	67	0.67
receiv	6	63	0.63
class	5	60	0.60
think	5	58	0.58
provid	6	56	0.56
school	6	56	0.56
institut	8	55	0.55
first	5	54	0.54
howev	5	54	0.54

degre	5	52	0.52
supervisor	10	52	0.52
correct	7	51	0.51
expect	6	51	0.51
tutor	5	50	0.50
undergradu	10	50	0.50

Table 65 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Development opportunities (including teaching / demonstrating)".

#### Overlap between references to *Teach* and *Student*

- Teaching is extremely beneficial and rewarding, however it takes up a lot of time and is always unpaid. Depending on supervisors, some students may have little to no support as they attempt to teach an entire module. Training is certainly necessary for this.
- As a part-time student, I couldn't attend the vast majority of the workshops provided for research students, as they are run during business hours. I would appreciate the chance to talk to someone about my career options. Although I am happy in my current career, one of the reasons I undertook this PhD was to give me the option of moving into academia. But I have no idea how to make myself more employable should I wish to pursue that route. I would love to have had the opportunity to teach, but being a part-time student, it would have been difficult to work around my full-time job.
- I was given no training for my considerable teaching responsibilities, and felt that support and guidance was very much lacking. Nor was I paid for any of the teaching that I did. While I very much enjoy teaching, the approach towards PhD students and Early Career Researchers, in which they seem to be used as free labour, is very disheartening.
- Teaching/demonstrating should be a paid activity. This is the case for most departments in \*\*\* (including my own department). However, I think it should be mentioned here that \*\*\* Physics department does not pay their tutors/demonstrators. It is simply not acceptable that PhD students should take on demonstration, tutoring and the associated corrections (which can take up to a full day) without pay while the lecturers responsible for these courses (who are getting paid) do little or no work.
- I did not get teaching work within my institution where I was a student but I did from a different third level institution and it was very helpful in improving my skills.

Table 66 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Development opportunities (including teaching / demonstrating)".

#### A.5.2 Matrix coding and statistical inference

Year of fieldwork	Development opportunities (including teaching / demonstrating)		Total
	Student	Teach	
2018	127 (37.1%)	215 (62.9%)	342 (100.0%)
2019	128 (40.1%)	191 (59.9%)	319 (100.0%)
Total	255 (38.6%)	406 (61.4%)	661 (100.0%)

Table 67 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Development opportunities (including teaching / demonstrating)".

For *Development opportunities (including teaching / demonstrating)*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Student* and *Teach* is not statistically significant ( $p = 0.4781$ ).

Programme type	Development opportunities (including teaching / demonstrating)		Total
	Student	Teach	
Masters	28 (38.4%)	45 (61.6%)	73 (100.0%)
PhD	227 (38.6%)	361 (61.4%)	588 (100.0%)
Total	255 (38.6%)	406 (61.4%)	661 (100.0%)

Table 68 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to “Development opportunities (including teaching / demonstrating)”.

For *Development opportunities (including teaching / demonstrating)* and the two most frequently occurring words – *Student* and *Teach* – the difference in distribution of references between programme types is not statistically significant ( $p = 1.000$ ).

## A.6 Research skills development

### A.6.1 Word frequency analysis



Figure 18 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Research skills development”.

Word	Length	Count	Weighted Percentage (%)
develop	7	80	3.54
Skill	5	80	3.54
Train	5	34	1.5
student	7	33	1.46
Integr	6	33	1.46
Modul	5	32	1.41
learn	5	30	1.33
cours	5	29	1.28
supervisor	10	29	1.28
ethic	5	28	1.24
institut	8	26	1.15
understand	10	25	1.1
univers	7	22	0.97
improv	6	22	0.97
programm	8	19	0.84
degre	5	17	0.75
confid	6	16	0.71
think	5	16	0.71
creativ	7	16	0.71
still	5	15	0.66
support	7	15	0.66
realli	6	15	0.66
provid	6	13	0.57
method	6	13	0.57
complet	7	13	0.57
appli	5	13	0.57
attend	6	13	0.57
experi	6	13	0.57
opportun	8	13	0.57
studi	5	12	0.53

Table 69 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Research skills development".

Overlap between references to *Develop/Skill\** and *Train*

- Research skills training courses within the \*\*\* have assisted in developing my knowledge of these. \*\*\* offers online 'research integrity' courses via our in-house IT platform.
- Access to a statistician specifically for research studies/PhD students is also lacking. The onus is left on the student to obtain statistical advice, whereas I feel this should be supported (in part at least) by the institute e.g. each school or department is assigned a statistician that they can access by appointment if required.
- My development of research skills has been as a result of training I have purchased outside of my institute not what is on offer withing the university.
- My Research skills developed from experience not formal training
- I certainly feel development. The research integrity training here seems like an obligation; the courses are poorly structured and timed, and don't supply much that couldn't be handled with a circulated document (aside from taking attendance...) My development in method and analysis has come through my research efforts, not through any provided training (which, to my knowledge, is slight)

Table 70 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Research skills development".

\* Referenced equally with same weighted percentage, see Table 69.

### A.6.2 Matrix coding and statistical inference

Year of fieldwork	Research skills development		Total
	Develop/Skill	Train	
2018	38 (79.2%)	10 (20.8%)	48 (100.0%)
2019	34 (68.0%)	16 (32.0%)	50 (100.0%)
Total	72 (73.5%)	26 (26.5%)	98 (100.0%)

Table 71 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Research skills development".

For *Research skills development*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Develop/Skill* and *Train* is not statistically significant ( $p = 0.3064$ ).

Programme type	Research skills development		Total
	Develop/Skill	Train	
Masters	8 (80.0%)	2 (20.0%)	10 (100.0%)
PhD	64 (72.7%)	24 (27.3%)	88 (100.0%)
Total	72 (73.5%)	26 (26.5%)	98 (100.0%)

Table 72 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to "Research skills development".



audienc	7	12	0.70
industri	8	11	0.64
realli	6	11	0.64
previous	8	11	0.64
train	5	11	0.64
learn	5	11	0.64
contact	7	10	0.59
studi	5	10	0.59
improv	6	10	0.59
person	6	10	0.59
alreadi	7	9	0.53
cours	5	9	0.53
increas	7	9	0.53

Table 73 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Development of other transferable skills”.

Overlap between references to <i>Develop</i> and <i>Skill</i>	
•	No formal emphasis is placed on any of the development of these transferrable skills. Therefore, any that have occurred are due to the requirements to progress with research. If any of these transferrable skills are hard to pick up, nothing is in place to monitor or assist with these.
•	Have not developed my writing skills as much, therefore the publishing process is still relatively unknown to me.
•	I drove my own development in transferable skills, I have taught modules abroad and in Ireland and I feel my communication skills and ability to deal with difficult personalities and speak in front of large audiences is greatly improved. My supervisor took no interest in my personal development and made no suggestions.
•	Skills and abilities have been developed outside of the programme not because of it.
•	Development of transferable skills, my professional network etc. has been mostly self-driven.

Table 74 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to “Development of other transferable skills”.

### A.7.2 Matrix coding and statistical inference

Year of fieldwork	Development of other transferable skills		Total
	Develop	Skill	
2018	37 (52.1%)	34 (47.9%)	71 (100.0%)
2019	38 (55.9%)	30 (44.1%)	68 (100.0%)
Total	75 (54.0%)	64 (46.0%)	139 (100.0%)

Table 75 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Development of other transferable skills”.

For *Development of other transferable skills*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Develop* and *Skill* is not statistically significant ( $p = 0.7829$ ).

Domicile group	Development of other transferable skills		Total
	Develop	Skill	
Irish	55 (53.4%)	48 (46.6%)	103 (100.0%)
Non-Irish	20 (55.6%)	16 (44.4%)	36 (100.0%)
Total	75 (54.0%)	64 (46.0%)	139 (100.0%)

Table 76 Matrix coding of the relationship between Domicile group and the two most frequently occurring words to the question relating to “Development of other transferable skills”.

For *Development of other transferable skills* and the two most frequently occurring words – *Develop* and *Skill* – the difference in distribution of references between domicile group is not statistically significant ( $p = 0.9766$ ).

## A.8 Student / staff responsibilities and supports

### A.8.1 Word frequency analysis



Figure 20 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Student / staff responsibilities and supports”.

Word	Length	Count	Weighted Percentage (%)
student	7	304	6.21
support	7	162	3.31
feedback	8	82	1.67
staff	5	75	1.53
supervisor	10	73	1.49
depart	6	59	1.20
respons	7	53	1.08
institut	8	45	0.92

school	6	42	0.86
univers	7	41	0.84
postgradu	9	40	0.82
avail	5	39	0.80
provid	6	31	0.63
health	6	31	0.63
think	5	31	0.63
counsel	7	30	0.61
chang	5	29	0.59
servic	6	28	0.57
inform	6	27	0.55
postgrad	8	26	0.53
regard	6	26	0.53

Table 77 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Student / staff responsibilities and supports".

#### Overlap between references to *Student* and *Support*

- I don't feel that supports, especially those in regards to graduate students are communicated appropriately. I think graduate students are some of the most vulnerable and often struggle with mental health issues and personal issues which can lead them to taking time away from their studies.
- Fees increase each year which puts considerable financial pressure on the self-funding student who has no financial support.
- Except my supervisor support and , I didn't receive any other information or support, contrary, I felt excluded. Moreover, they asked me help when they needed it answers for a survey made by another student, using abusively my email address.
- The university needs to be more proactive in terms of supporting PhD students and creating healthy work environments. The mental health of many PhD students is extremely poor.
- I know of various student supports, however I would not usually have the need to seek them out. If/when I do, I know where to go in order to get the information that I need.

Table 78 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Student / staff responsibilities and supports".

### A.8.2 Matrix coding and statistical inference

Year of fieldwork	Student / staff responsibilities and supports		Total
	Student	Support	
2018	101 (60.5%)	66 (39.5%)	167 (100.0%)
2019	95 (62.1%)	58 (37.9%)	153 (100.0%)
Total	196 (61.3%)	124 (38.8%)	320 (100.0%)

Table 79 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Student / staff responsibilities and supports".

For *Student / staff responsibilities and supports*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Student* and *Support* is not statistically significant ( $p = 0.8564$ ).

Mode of study	Student / staff responsibilities and supports		Total
	Student	Support	
Full-Time	171 (62.2%)	104 (37.8%)	275 (100.0%)
Part-time/Remote	25 (55.6%)	20 (44.4%)	45 (100.0%)
Total	196 (61.3%)	124 (38.8%)	320 (100.0%)

Table 80 Matrix coding of the relationship between Mode of study and the two most frequently occurring words to the question relating to “Student / staff responsibilities and supports”.

For *Student / staff responsibilities and supports* and the two most frequently occurring words – *Student* and *Support* – the difference in distribution of references between mode of study is not statistically significant ( $p = 0.4960$ ).

## A.9 Personal outlook

### A.9.1 Word frequency analysis

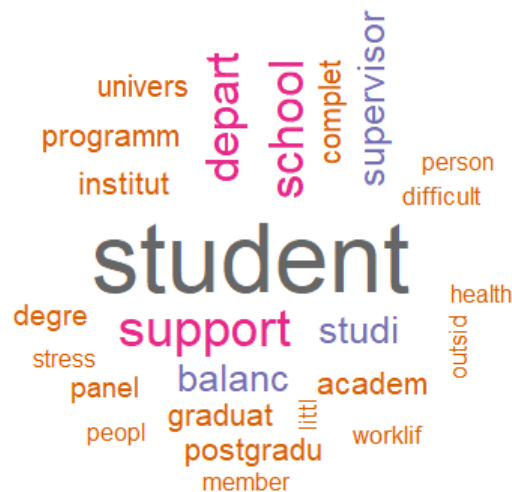


Figure 21 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Personal outlook”.

Word	Length	Count	Weighted Percentage (%)
student	7	117	3.54
support	7	55	1.66
school	6	49	1.48
depart	6	48	1.45
balanc	6	37	1.12
studi	5	37	1.12
supervisor	10	35	1.06
programm	8	29	0.88
institut	8	29	0.88
graduac	7	28	0.85
postgradu	9	27	0.82
academ	6	27	0.82
univers	7	26	0.79
complet	7	25	0.76
degre	5	23	0.70
panel	5	22	0.67
worklif	7	20	0.60
difficult	9	20	0.60
littl	5	19	0.57
outsid	6	18	0.54
person	6	18	0.54
member	6	17	0.51
stress	6	17	0.51
peopl	5	17	0.51
health	6	17	0.51

*Table 81 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Personal outlook".*

**Overlap between references to *Student* and *Support***

- While I enjoy my research degree programme, I spend a lot of time worrying about what career path it will bring me down and if it will be of use to me in my life after I graduate. I think a lot of students may feel this way and so some extra support regarding these problems might be beneficial to all students
- University should put more efforts to support research students rather than focusing on taught programmes.
- Student supports are inadequate. I was turned away from supports because I was not perceived as being in danger of failing to complete my degree programme, regardless of my personal outlook.
- I regret starting this program and there are little to no supports for the mental wellbeing of students. Nobody has ever come down to my study area and asked if everything is ok or if I am struggling with anything in my research, even when I have flagged that I am
- I have amazing fellow postgraduate students that I can share my problems with, however, because of the lack of support for research students I am not happy in my institution. I do not know if my research programme was worthwhile anymore as I see some of the recent graduates from my institute struggle to find work, especially in research. The majority of the students find jobs in quality control where a research degree is of very little value and not even required. It feels to me that a lot of job possibilities will not be available to me in the future because of the institution that I chose to do my degree in.

Table 82 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Personal outlook".

**A.9.2 Matrix coding and statistical inference**

Year of fieldwork	Personal outlook		Total
	Student	Support	
2018	0 (0.0%)	0 (0.0%)	0 (0.0%)
2019	77 (64.2%)	43 (35.8%)	120 (100.0%)
Total	77 (64.2%)	43 (35.8%)	120 (100.0%)

Table 83 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Personal outlook".

For *Personal outlook*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Student* and *Support* could not be determined due to there being no references in 2018.

Programme type	Personal outlook		Total
	Student	Support	
Masters	8 (61.5%)	5 (38.5%)	13 (100.0%)
PhD	69 (64.5%)	38 (35.5%)	107 (100.0%)
Total	77 (64.2%)	43 (35.8%)	120 (100.0%)

Table 84 Matrix coding of the relationship between Programme type and the two most frequently occurring words to the question relating to "Personal outlook".

For *Personal outlook* and the two most frequently occurring words – *Student* and *Support* – the difference in distribution of references across programme types is not statistically significant ( $p = 1.000$ ).

## A.10 Aspects / elements of research degree programme are most valuable

### A.10.1 Word frequency analysis

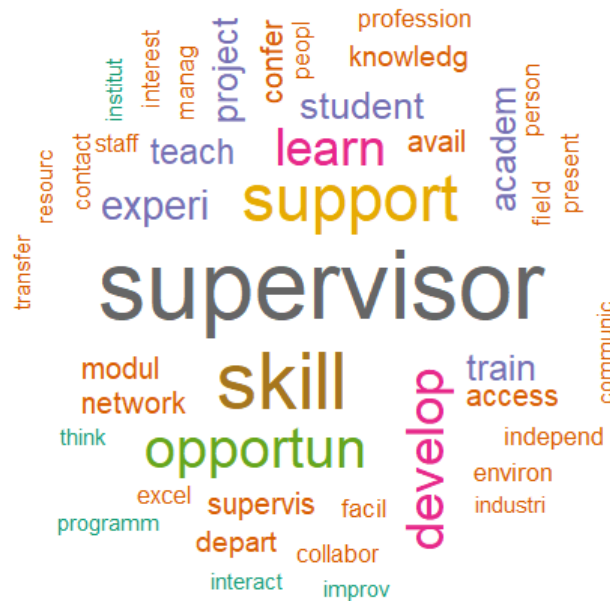


Figure 22 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Aspects / elements of research degree programme are most valuable”.

Word	Length	Count	Weighted Percentage (%)
supervisor	10	646	4.35
Skill	5	520	3.50
support	7	407	2.74
opportun	8	334	2.25
develop	7	314	2.11
Learn	5	313	2.11
Experi	6	218	1.47
student	7	197	1.33
train	5	183	1.23
academ	6	178	1.20
project	7	175	1.18
teach	5	174	1.17
modul	5	143	0.96
avail	5	142	0.96
network	7	140	0.94
supervis	8	139	0.94

access	6	137	0.92
confer	6	131	0.88
knowledg	8	120	0.81
depart	6	116	0.78
facil	5	109	0.73
environ	7	103	0.69
field	5	102	0.69
independ	8	102	0.69
profession	10	101	0.68
manag	5	101	0.68
collabor	8	95	0.64
present	7	95	0.64
staff	5	92	0.62
excel	5	91	0.61
person	6	90	0.61
peopl	5	90	0.61
interest	8	85	0.57
contact	7	85	0.57
communic	8	83	0.56
resourc	7	83	0.56
industri	8	83	0.56
transfer	8	81	0.55
think	5	79	0.53
programm	8	79	0.53
interact	8	75	0.50
institut	8	74	0.50
improv	6	74	0.50

Table 85 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to "Aspects / elements of research degree programme are most valuable".

#### Overlap between references to *Supervisor* and *Skill*

- the resources are available I have enough guidance from my supervisor but in the same time I have the time to work on my own and develop various skills such as time management, research skills , etc
- I am fortunate to have a fantastic supervisor who offers a lot of support with my research. There is also lots of opportunities to enhance transferable skills through Graduate studies.
- Getting along great with all supervisors as this is key for a good experience. But also a keen interest in the project as well as the opportunity to gain new skills.
- There is great support and training in my institution. My supervisor has also been great at helping me work towards a career in academia as well as ensuring I have the skills to complete my doctorate.
- That I have a supervisor who cares The wide range analysis skills that I'm learning

Table 86 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to "Aspects / elements of research degree programme are most valuable".

### A.10.2 Matrix coding and statistical inference

Year of fieldwork	Aspects / elements of research degree programme are most valuable		Total
	Skill	Supervisor	
2018	253 (41.9%)	351 (58.1%)	604 (100.0%)
2019	208 (41.1%)	298 (58.9%)	506 (100.0%)
Total	461 (41.5%)	649 (58.5%)	1110 (100.0%)

Table 87 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to "Aspects / elements of research degree programme are most valuable".

For *Aspects / elements of research degree programme are most valuable*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Skill* and *Supervisor* is not statistically significant ( $p = 0.8401$ ).

Gender	Aspects / elements of research degree programme are most valuable		Total
	Skill	Supervisor	
Female	304 (41.6%)	427 (58.4%)	731 (100.0%)
Male	157 (41.4%)	222 (58.6%)	379 (100.0%)
Total	461 (41.5%)	649 (58.5%)	1110 (100.0%)

Table 88 Matrix coding of the relationship between Gender and the two most frequently occurring words to the question relating to "Aspects / elements of research degree programme are most valuable"

For *Aspects / elements of research degree programme are most valuable* and the two most frequently occurring words – *Skill* and *Supervisor* – the difference in distribution of references with respect to gender is not statistically significant ( $p = 1.000$ ).

## A.11 Aspects of research degree experience could be improved

### A.11.1 Word frequency analysis

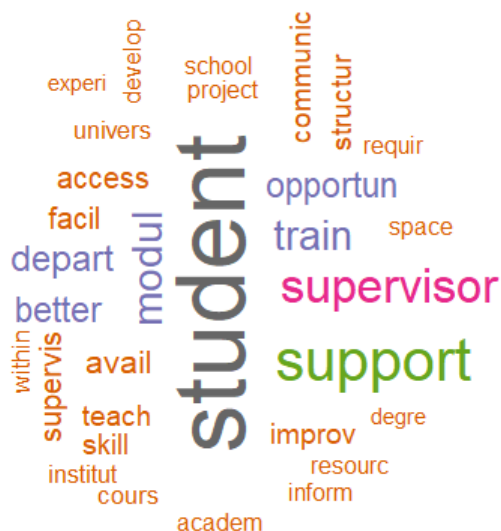


Figure 23 Word cloud representing the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Aspects of research degree experience could be improved”.

Word	Length	Count	Weighted Percentage (%)
student	7	815	3.90
support	7	468	2.24
supervisor	10	353	1.69
modul	5	285	1.36
Train	5	269	1.29
depart	6	224	1.07
Better	6	213	1.02
opportun	8	207	0.99
avail	5	188	0.90
access	6	180	0.86
supervis	8	170	0.81
facil	5	169	0.81
teach	5	163	0.78
improv	6	163	0.78
skill	5	159	0.76
structur	8	151	0.72
communic	8	150	0.72
univers	7	136	0.65
institut	8	131	0.63
cours	5	130	0.62
academ	6	123	0.59
inform	6	120	0.57
space	5	117	0.56

within	6	111	0.53
project	7	111	0.53
requir	6	109	0.52
resourc	7	109	0.52
school	6	109	0.52
develop	7	107	0.51
degre	5	106	0.51
experi	6	106	0.51

Table 89 RStudio word frequency analysis of the most frequently occurring words, with a weighted percentage of at least 0.5% to the question relating to “Aspects of research degree experience could be improved”.

Overlap between references to <i>Student</i> and <i>Support</i>
<ul style="list-style-type: none"> <li>• More facilities for a work-space where I can type but also listen (even through headphones) to the music I need to analyse. More visible supports for students also.</li> <li>• Student support, facilities, circulation of key information.</li> <li>• Scholarship student support services for easy spending. The bureaucracy of the university makes it very difficult to spend our little money.</li> <li>• Support for interdisciplinary nature of my research. Better networking opportunities with other research students in musicology.</li> <li>• Lack of student engagement across campuses. Little financial support for research activities in the Humanities.</li> </ul>

Table 90 Sample of reference extracts from the overlap of the two most frequently occurring words to the question relating to “Aspects of research degree experience could be improved”.

### A.11.2 Matrix coding and statistical inference

Year of fieldwork	Aspects of research degree experience could be improved		Total
	Student	Support	
2018	303 (55.7%)	241 (44.3%)	544 (100.0%)
2019	291 (63.1%)	170 (36.9%)	461 (100.0%)
Total	594 (59.1%)	411 (40.9%)	1005 (100.0%)

Table 91 Matrix coding of the relationship between Year of fieldwork and the two most frequently occurring words to the question relating to “Aspects of research degree experience could be improved”.

For *Aspects / elements of research degree programme are most valuable*, from 2018 to 2019, the change in distribution of student responses referencing stemmed variants of *Student* and *Support* is statistically significant ( $p = 0.0203$ ). In 2019, there are more references to *Student* and less to *Support* than in 2018.

HEI type	Aspects of research degree experience could be improved		Total
	Student	Support	
Institutes of Technology	106 (60.2%)	70 (39.8%)	176 (100.0%)
Universities	473 (58.9%)	330 (41.1%)	803 (100.0%)
Other Institutions	15 (57.7%)	11 (42.3%)	26 (100.0%)
Total	594 (59.1%)	411 (40.9%)	1005 (100.0%)

Table 92 Matrix coding of the relationship between HEI type and the two most frequently occurring words to the question relating to "Aspects of research degree experience could be improved".

For *Aspects of research degree experience could be improved* and the two most frequently occurring words – *Student* and *Support* – the difference in distribution of references with respect to HEI type is not statistically significant ( $p = 0.9387$ ).

Note: Although there is a statistically significant difference in reference distribution from 2018 to 2019, the difference did not impact reference distribution between HEI types from 2018 – 2019 – i.e., the difference in reference distribution in 2018 and 2019 across HEI types were not statistically significant, hence, results presented in Table 92 are not split by year.

## A.12 Required RStudio packages

*library(readxl) # import data from Excel*

*library(data.table) # rename variables*

*library(tm) # required for text mining*

*library(SnowballC) # for text stemming (reduces words to their stemmed core)*

*library(wordcloud) # word-cloud generator*

*library(RColorBrewer) # colour palettes in word-cloud*

*library(dplyr) # required for filtering*

*library(xlsx) # write results to Excel*

*library(stringr) # reference extracting*

*library(tidyr) # formatting dataframes*

*library(summarytools) # for matrix coding / crosstabulations*